



PT diagnosis, prevention and treatment of disorders associated with  
PT abnormal nucleolin-like protein expression such as cancers,  
XX autoimmune disorders and Alzheimer's disease

PS Claim 8; Fig 1; 33pp; English.

CC The present sequence represents human nucleolin-like peptide, designated  
CC HNLP, which is the main protein component in the nucleolus of eukaryotic  
CC cells and is an essential part of ribosome biosynthesis and also plays  
CC an important role in importing proteins to the nucleus. HNLP may be used  
CC in the diagnosis, prevention and treatment of disorders associated with  
CC abnormal expression of HNLP. For example, it may be used to treat  
CC cancers (e.g. melanoma, breast cancer and prostate cancer), autoimmune  
CC disorders (e.g. autoimmune haemolytic anaemia and inflammatory bowel  
CC disease) and, in particular, Alzheimer's disease. It may also be used to  
CC study the function of the HNLP peptides, the formation (biosynthesis) of  
CC ribosomes and the intake of proteins into the nucleus.

XX Sequence 545 AA;

Query Match 100.0%; Score 2898; DB 20; Length 545;  
Best Local Similarity 100.0%; Pred. No. 7.6e-243;  
Matches 545; Conservative 0; Indels 0; Gaps 0;

Db 1 MATIEVNGNCTEEPMDTTAVTHSENFCQLDAGLPOKVAEKLDEIYNYGLVAHSQDLER 60

Db 1 MATIEVNGNCTEEPMDTTAVTHSENFCQLDAGLPOKVAEKLDEIYNYGLVAHSQDLER 60

Qy 61 AIEALKREFINEDGALVQQLPFDSDLSHVQNSKAFLCGVNKTYRQEKGTKVADSSKGDF 120

Db 61 AIEALKREFINEDGALVQQLPFDSDLSHVQNSKAFLCGVNKTYRQEKGTKVADSSKGDF 120

Qy 121 EAKTAKLLERITYGLTDVTTGORKYTGPPDSVYSGQQPSVGTEIFVGKIPRDLFPEDELVP 180

Db 121 EAKTAKLLERITYGLTDVTTGORKYTGPPDSVYSGQQPSVGTEIFVGKIPRDLFPEDELVP 180

Qy 181 LFEKAGPIMDLRLMMDPDPLTGLNRGTAFTRCTKEAAQAVKLNNHEITSKGHIGVCTSV 240

Db 181 LFEKAGPIMDLRLMMDPDPLTGLNRGTAFTRCTKEAAQAVKLNNHEITSKGHIGVCTSV 240

Qy 241 ANNRLFGSIPKSITKMEQILJBEFSKVTEGLTDVLYHQDDKKNRGFPLETEDHKTA 300

Db 241 ANNRLFGSIPKSITKMEQILJBEFSKVTEGLTDVLYHQDDKKNRGFPLETEDHKTA 300

Qy 301 QARRRLMSGVKUWGNVGTWADPLDPREMAVKVYLFRNLANTVTEELKEAFSQ 360

Db 301 QARRRLMSGVKUWGNVGTWADPLDPREMAVKVYLFRNLANTVTEELKEAFSQ 360

Qy 361 FGKLERVKLKDYAPTHEDRGAVKAMBEMNGKDLLEGEMIEIVFAKPPDQKRKERKACR 420

Db 361 FGKLERVKLKDYAPTHEDRGAVKAMBEMNGKDLLEGEMIEIVFAKPPDQKRKERKACR 420

Qy 421 QAQRNQMYDDYYGGPPHMPPTRGRGRGRRGGYGYPPDYDYYGYDHNYRG 480

Db 421 QAQRNQMYDDYYGGPPHMPPTRGRGRGRRGGYGYPPYYGYDYYGYDHNYRG 480

Qy 481 YEDPYGYEDFOVGARGRGRGARGAAAPSRRGGAAPPGRAGYSORGPGSARGVRAGR 540

Db 481 YEDPYGYEDFOVGARGRGRGARGAAAPSRRGGAAPPGRAGYSORGPGSARGVRAGR 540

Qy 541 GRGRS 545

Db 541 GRGRS 545

Qy 100.0%; Score 2898; DB 23; Length 545;

Best Local Similarity 100.0%; Pred. No. 7.6e-243;

Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MATEHVNGNGTEBMDTTSAVIHSENFOTLIDAGLPOKVAEKLDEIYAGLVRAHSQDLER 60

DT 1 MATEHVNGNGTEBMDTTSAVIHSENFOTLIDAGLPOKVAEKLDEIYAGLVRAHSQDLER 60

PT Human nucleolin-like protein, HNLP.

PT Human: nucleolin-like protein; HNLP: autoimmune disorder;  
XX acquired immunodeficiency syndrome; AIDS; Addison's disease; allergy;  
KW asthma; atherosclerosis; multiple sclerosis; rheumatoid arthritis;  
KW osteoporosis; viral infection; bacterial infection; fungal infection;  
KW parasitic infection; protozoal infection; helminthic infection; cancer;  
KW Alzheimer's disease; systemic sclerosis; graft-versus-host disease;  
KW systemic lupus erythematosus; INCYTE 280975.

CC Homo sapiens.

XX US2002098566-A1.

PD 25-JUL-2002.

XX 15-OCT-2001; 2001US-0978242.

XX PR 12-DEC-1997; 97US-0990114.

XX PR 01-FEB-1999; 99US-0241333.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Bandman O, Yue H, Corley NC, Shah P;

XX DR 2002-690482/74.

XX DR N-PSDB; ABSS299.

XX Novel human nucleolin-like polypeptide, useful in diagnosis, prevention  
PT and treatment of cancer, Alzheimer's disease and autoimmune disorder  
PT such as AIDS, Addison's disease, allergy, asthma, and atherosclerosis -  
XX

CC Claim 1: Fig 1: 37PP: English.

XX The invention relates to an isolated human nucleolin-like polypeptide  
CC (HNLP) (S1), a polypeptide comprising a naturally occurring sequence  
CC having at least 90% identity to S1, or a biologically active or  
CC immunogenic fragment of S1 and the HNLP encoding nucleic acid. HNLP is  
CC useful for screening a compound for effectiveness as an agonist or  
CC antagonist, for screening a compound that specifically binds HNLP or  
CC modulates the activity of HNLP, and for preparing a polyclonal or  
CC monoclonal antibody by hybridoma technology. HNLP nucleic acid is useful  
CC for screening a compound for effectiveness in altering expression of a  
CC target polynucleotide comprising HNLP nucleic acid and HNLP probes are  
CC useful for assessing toxicity of a test compound. Anti-HNLP antibody is  
CC useful in a diagnostic test for a condition or a disease associated with  
CC the expression of HNLP in a biological sample, for detecting HNLP in a  
CC sample, and for preparing HNLP from a sample. HNLP ant agonists are  
CC useful for treating a disease or condition associated with decreased or  
CC increased expression of functional HNLP. The antibody is useful  
CC for diagnosing a condition or disease associated with the expression of  
CC HNLP in a subject. A HNLP nucleic acid microarray is useful for  
CC generating a transcript image of a sample which contains polynucleotides.  
CC HNLP and its nucleic acid are useful for diagnosing, treating and  
CC preventing an autoimmune disorder (e.g. acquired immunodeficiency  
CC syndrome (AIDS), Alzheimer's disease, allergy, asthma, atherosclerosis,  
CC multiple sclerosis, rheumatoid arthritis), osteoporosis, viral, bacterial,  
CC fungal, parasitic, protozoal, helminthic infections, cancer,  
CC Alzheimer's disease, systemic sclerosis, graft-versus-host disease and  
CC systemic lupus erythematosus (many more diseases are listed in the  
CC specification). The present sequence is the human HNLP protein encoded by  
CC a cDNA from INCYTE clone 280975.

XX SQ Sequence 545 AA;

XX Query Match 100.0%; Score 2898; DB 23; Length 545;

Best Local Similarity 100.0%; Pred. No. 7.6e-243;

Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MATEHVNGNGTEBMDTTSAVIHSENFOTLIDAGLPOKVAEKLDEIYAGLVRAHSQDLER 60

Db 1 MATEHVNGNGTEBMDTTSAVIHSENFOTLIDAGLPOKVAEKLDEIYAGLVRAHSQDLER 60

Best Local Similarity 100.0%; Pred. No. 6.2e-239; Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FILING DATE: ; ATTORNEY/AGENT INFORMATION: ; BILLING, Lucy J.; REGISTRATION NUMBER: 36,749; REFERENCE/DOCKET NUMBER: PF-0451 US; TELEPHONE: ; TELEFAX: 650-855-4166; TELEX: ; INFORMATION FOR SEQ ID NO: 1: ; SEQUENCE CHARACTERISTICS: ; LENGTH: 545 amino acids; TYPE: amino acid; STRANDEDNESS: single; TOPOLOGY: linear; IMMEDIATE SOURCE: ; LIBRARY: TLYMNOT05; CLONE: 2809795; US-09-241-333-1:

Query Match 100.0%; Score 2898; Db 4; Length 545; Best Local Similarity 100.0%; Pred. No. 6.2e-239; Mismatches 0; Indels 0; Gaps 0; Matches 545; Conservative 0;

Qy 1 MATHEVNGNTGTEPMOTTSAVTHSENQTLDAGLPQKVAELDEIYAGLVAHSDLDER 60  
Db 1 MATHEVNGNTGTEPMOTTSAVTHSENQTLDAGLPQKVAELDEIYAGLVAHSDLDER 60  
Qy 61 AIEALKSFNEQDGAALVLIQFKQFDSDLSHQNKSFLCGVMKTYTREPKQGTKVADSSKGPD 120  
Db 61 AIEALKSFNEQDGAALVLIQFKQFDSDLSHQNKSFLCGVMKTYTREPKQGTKVADSSKGPD 120  
Qy 121 EAKIKALLERTGYTLDVTGORKYGGPPDSVYSGQQPSVSTEIFVGKIPRDLFEDLYP 180  
Db 121 EAKIKALLERTGYTLDVTGORKYGGPPDSVYSGQQPSVSTEIFVGKIPRDLFEDLYP 180  
Qy 181 LPEKAGPTWDLRMMPLTGRLNRGYAFVTFCTKEAAQEAVLYNNHEIRSGKHHIGVCISV 240  
Db 181 LPEKAGPTWDLRMMPLTGRLNRGYAFVTFCTKEAAQEAVLYNNHEIRSGKHHIGVCISV 240  
Qy 241 ANNRLFVGSPISPKSKTKBQIILEFSKTYEGLTDVILYHQPDDKKRNQFCPEYEDHKTA 300  
Db 241 ANNRLFVGSPISPKSKTKBQIILEFSKTYEGLTDVILYHQPDDKKRNQFCPEYEDHKTA 300  
Qy 301 QARRRLMSGKYKVGNGTVBWDPLIEDPDEVMAKVVKLFVRNLANTYTEELEKAFFSQ 360  
Db 301 QARRRLMSGKYKVGNGTVBWDPLIEDPDEVMAKVVKLFVRNLANTYTEELEKAFFSQ 360  
Qy 361 FGKLERVKLKDYAIFHFDERDGAVKAMEEENKGDKLEGENTIEIVPAKPDQKRKERAQK 420  
Db 361 FGKLERVKLKDYAIFHFDERDGAVKAMEEENKGDKLEGENTIEIVPAKPDQKRKERAQK 420  
Qy 421 QAAKNQMDYYGGPHMPPTGRGRGGCYGPPYYGYEDYYGYDHYNYRG 480  
Db 421 QAAKNQMDYYGGPHMPPTGRGRGGCYGPPYYGYEDYYGYDHYNYRG 480  
Qy 481 YEDPYGYEDFQVARGRRGARGAAPSRRGAAAPPGRAGYSORGGSAREVRAGR 540  
Db 481 YEDPYGYEDFQVARGRRGARGAAPSRRGAAAPPGRAGYSORGGSARGVRAGR 540  
Qy 541 GRGRS 545  
Db 541 GRGRS 545

RESULT 2  
US-09-241-333-1  
Sequence 1, Application US/09241333  
Patent No. 6313266  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Yue, Henry  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/241,333  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/990,114

RESULT 3  
US-09-347-833-2  
Sequence 2, Application US/09347833  
Patent No. 6394658  
GENERAL INFORMATION:  
APPLICANT: Famodu, Layo O.  
APPLICANT: Odell, Joan T.  
TITLE OF INVENTION: Factors Involved in Gene Expression

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 09:23:45 ; Search time 21 Seconds  
(without alignments)  
1098.068 Million cell updates/sec

Title: US-09-978-242-1  
Perfect score: 2898  
Sequence: 1 MATEHVNNGNCTTEPMDDTSA.....RGGPGSARGVRAGKRGGRS 545

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
 1: /cgn2\_6.ptodata/1/iaa/5A\_COMB.pep:/\*  
 2: /cgn2\_6.ptodata/1/iaa/5B\_COMB.pep:/\*  
 3: /cgn2\_6.ptodata/1/iaa/6A\_COMB.pep:/\*  
 4: /cgn2\_6.ptodata/1/iaa/6B\_COMB.pep:/\*  
 5: /cgn2\_6.ptodata/1/iaa/BctUS\_COMB.pep:/\*  
 6: /cgn2\_6.ptodata/1/iaa/backFiles1.pep:/\*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**RESULTS**

RESULT 1  
US-09-990-114-1  
; Sequence 1, Application US/08990114  
; Patent No. 592475

; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Yue, Henry  
 ; APPLICANT: Corley, Neil C.  
 ; APPLICANT: Shah, Purvi  
 ; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ For Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/990,114  
 ; FILING DATE: Herewith  
 ; CLASSIFICATION:  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0451 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-855-0555  
 ; TELEFAX: 650-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 545 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: TLYMNNT05  
 ; CLONE: 2806795  
 ; US-08-990-114-1

Query Match 100.0%; Score 2898; DB 2; Length 545;

Result No.	Score	Query	Match	Length	DB ID	Description
1	2898	100.0	545	2	US-08-990-114-1	Sequence 1, Appli
2	2898	100.0	545	4	US-09-241-333-1	Sequence 1, Appli
3	299.5	10.3	652	3	US-09-347-833-2	Sequence 2, Appli
4	269	9.3	655	3	US-09-347-833-4	Sequence 4, Appli
5	258.5	8.9	714	2	US-08-990-114-3	Sequence 3, Appli
6	258.5	8.9	714	4	US-09-241-333-3	Sequence 3, Appli
7	223.5	7.7	336	1	US-07-667-276A-8	Sequence 8, Appli
8	209	7.2	359	1	US-07-881-015-2	Sequence 2, Appli
9	209	7.2	359	1	US-08-120-827-2	Sequence 2, Appli
10	209	7.2	359	1	US-08-147-675-2	Sequence 2, Appli
11	206.5	7.1	380	1	US-07-881-015-51	Sequence 51, Appli
12	206.5	7.1	380	1	US-08-120-827-51	Sequence 51, Appli
13	206.5	7.1	380	1	US-08-478-675-51	Sequence 5, Appli
14	205.5	7.1	675	3	US-08-573-273-5	Sequence 5, Appli
15	204.5	7.1	688	3	US-08-573-273-26	Sequence 26, Appli
16	204.5	7.1	747	3	US-08-973-273-3	Sequence 3, Appli
17	197.5	6.8	341	2	US-08-538-711A-8	Sequence 8, Appli
18	197.5	6.8	341	3	US-08-725-027-8	Sequence 8, Appli
19	197.5	6.8	341	4	US-09-542-552-8	Sequence 8, Appli
20	197.5	6.8	353	2	US-08-538-711A-7	Sequence 7, Appli
21	197.5	6.8	353	4	US-09-542-552-7	Sequence 7, Appli
22	197.5	6.8	444	1	US-07-881-075-3	Sequence 3, Appli
23	197.5	6.8	444	1	US-08-120-827-3	Sequence 3, Appli
24	197.5	6.8	444	1	US-08-478-675-3	Sequence 3, Appli
25	194.5	6.7	223	1	US-07-667-276A-10	Sequence 10, Appli
26	194.5	6.7	428	3	US-09-347-833-6	Sequence 6, Appli

TELEFAX: 650-845-4166  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 545 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: TLYMNOTOS  
 CLONE: 2899795  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-978-242-1

Query Match 100.0%; Score 2098; DB 9; Length 545;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-239; Indels 0; Gaps 0;  
 Matches 545; Conservative 0; Mismatches 2; Gaps 1;

Qy 1 MATHEVNGNTTEEPMDTTSAVHSNFQTLLDAGLPQKVAEKLDEIYAGLVAHSDLDER 60  
 Db 1 MATHEVNGNTTEEPMDTTSAVHSNFQTLLDAGLPQKVAEKLDEIYAGLVAHSDLDER 60  
 Qy 61 AIEALKEFNSDGLAVLQOQFKDSLISHVQNKSAFLCGVMKTYREKEQTKVADSSKGPD 120  
 Db 61 AIEALKEFNSDGLAVLQOQFKDSLISHVQNKSAFLCGVMKTYREKEQTKVADSSKGPD 120  
 Qy 121 EAKIKALLBERTGTYTLDTTGQRKYGGPPDSTSQQGQPSVGTEIVFGKIPRDLFEDELVP 180  
 Db 121 EAKIKALLBERTGTYTLDTTGQRKYGGPPDSTSQQGQPSVGTEIVFGKIPRDLFEDELVP 180  
 Qy 122 LFEKAGPIMDLRLLMDPLTGLNRGYAFVTPCTKEAQAEEVAKLYNNHEITRSKGHIGVCISV 240  
 Db 122 LFEKAGPIMDLRLLMDPLTGLNRGYAFVTPCTKEAQAEEVAKLYNNHEITRSKGHIGVCISV 240  
 Qy 181 LFEKAGPIMDLRLLMDPLTGLNRGYAFVTPCTKEAQAEEVAKLYNNHEITRSKGHIGVCISV 240  
 Db 181 LFEKAGPIMDLRLLMDPLTGLNRGYAFVTPCTKEAQAEEVAKLYNNHEITRSKGHIGVCISV 240  
 Qy 241 ANNRLFVGSIPKSXKTEQILEEFSKTEGLDVKKRNRCFCFLEYEDHTAA 300  
 Db 241 ANNRLFVGSIPKSXKTEQILEEFSKTEGLDVKKRNRCFCFLEYEDHTAA 300  
 Qy 258 ANNRLFVGSIPKSXKTEQILEEFSKTEGLDVKKRNRCFCFLEYEDHTAA 317  
 Db 258 ANNRLFVGSIPKSXKTEQILEEFSKTEGLDVKKRNRCFCFLEYEDHTAA 317  
 Qy 301 QAARRLMSCKVKWVNGNTVEADPIDEPPVEMAKYKVLFVNLANTVTEILEKAFAQSO 360  
 Db 301 QAARRLMSCKVKWVNGNTVEADPIDEPPVEMAKYKVLFVNLANTVTEILEKAFAQSO 360  
 Qy 318 QAARRLMSCKVKWVNGNTVEADPIDEPPVEMAKYKVLFVNLANTVTEILEKAFAQSO 377  
 Db 318 QAARRLMSCKVKWVNGNTVEADPIDEPPVEMAKYKVLFVNLANTVTEILEKAFAQSO 377  
 Qy 361 FGKLERVKLKDYAFHDERGAVKAMBEMNGKDLEGENTIEVAKPDPDKRKERKAQR 420  
 Db 361 FGKLERVKLKDYAFHDERGAVKAMBEMNGKDLEGENTIEVAKPDPDKRKERKAQR 420  
 Qy 378 FGKLERVKLKDYAFHDERGAVKAMBEMNGKDLEGENTIEVAKPDPDKRKERKAQR 437  
 Db 378 FGKLERVKLKDYAFHDERGAVKAMBEMNGKDLEGENTIEVAKPDPDKRKERKAQR 437  
 Qy 421 QAAKQMDDYYYYGPAMPMPPTGRGRGRRGGYGPDDYYGYEDYYGYDYYHNYRG 480  
 Db 421 QAAKQMDDYYYYGPAMPMPPTGRGRGRRGGYGPDDYYGYEDYYGYDYYHNYRG 480  
 Qy 438 QAAKQMDDYYYYGPMPMPPTGRGRGRRGGYGPDDYYGYEDYYGYDYYHNYTRGG 497  
 Db 438 QAAKQMDDYYYYGPMPMPPTGRGRGRRGGYGPDDYYGYEDYYGYDYYHNYTRGG 497  
 Qy 481 YEDPYGYEDFOVQARGRRGRRGARGAAPSRGRAAAPPRGRAGYSORGPGSARGVRGARG 539  
 Db 481 YEDPYGYEDFOVQARGRRGRRGARGAAPSRGRAAAPPRGRAGYSORGPGSARGVRGARG 539  
 Qy 498 YEDPYGYEDFOVQARGRRGRRGARGAAPSRGRAAAPPRGRAGYSORGPGSARGVRGARG 557  
 Db 498 YEDPYGYEDFOVQARGRRGRRGARGAAPSRGRAAAPPRGRAGYSORGPGSARGVRGARG 557  
 Qy 540 ---RGRGR 544  
 Db 540 ---RGRGR 544  
 Qy 558 GAQQQRGRSQ 567  
 Db 558 GAQQQRGRSQ 567

RESULT 3  
 US-10-106-698-5319  
 ; Sequence 5319, Application US/10106698  
 ; Publication No. US20030109690A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide;  
 ; FILE REFERENCE: PA005P1  
 ; CURRENT APPLICATION NUMBER: US/10/106-698  
 ; PRIORITY INFORMATION:  
 ; CURRENT FILING DATE: 2002-03-27  
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524  
 ; PRIOR FILING DATE: 2000-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/157,137  
 ; PRIOR FILING DATE: 1999-09-29  
 ; PRIOR APPLICATION NUMBER: US 60/163,280  
 ; PRIOR FILING DATE: 1999-11-03  
 ; NUMBER OF SEQ ID NOS: 8554  
 ; SOFTWARE: PatentIn Ver. 3.0  
 ; SEQ ID NO 5319

RESULT 2  
 US-09-925-300-1415  
 ; Sequence 1415, Application US/09925300  
 ; Patent No. US2002015168A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben,  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/0598  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270



Qy 474 YHNYRGGYEDPYGYED-FOVGARGRGGRGARGA-APSRGRGAAAPRGRAGYSGORGPP G 530  
 Db 479 YHDYRGGYEPYGYDDGAVGRG-GRGRGGRGAPPGRGAGYSGORGAPG 537

RESULT 3  
 hypothetical protein A-TM018A10.14 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Accession: T01563  
 R;Dempsey, S.; Harper, M.  
 Submitted to the EMBL Data Library, July 1997  
 A;Description: The sequence of A. thaliana TM018A10.  
 A;Reference number: Z14348  
 A;Map position: 4  
 A;Status: translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Residues: 1-521 <DEML>  
 A;Cross-references: EMBL:AF013294; NID:92252848; PID:92252863  
 A;Experimental source: Cultivar Columbia  
 C;Genetics:  
 A;Map position: 4  
 A;Introns: 158/3; 209/3; 256/3; 316/3; 384/1; 444/3; 473/2  
 A;Note: A\_TM018A10.14

Query Match Score 495.5; DB 2; Length 521;  
 Best Local Similarity 26.2%; Pred. No. 3.9e-29;  
 Matches 149; Conservative 97; Mismatches 178; Indels 145; Gaps 17;

Qy 53 AHSDL-DERAIALKEFNEFGCALAVLQQPKFDSDLSHVQNSKAFLCGVMMTYQRBEKQGTK 111  
 Db 17 SYSEMDDEVEEQVEEYE-----EEEDDDDDVGQNA-----BEREVD 58

Qy 112 VADSSKGPDPEAKIKAIALLERTSYTLDVTGORKYKGPPP----DSVYSG- -QQPSSVGTI 164  
 Db 59 YGDTRGGDMEDVQEETAEDDNHIDLETADDEKFPSPIIDREKYSVHSLSLPHGSEV 118

Qy 165 FVGKIPRLDFDELVLPLFEXAKPIWD-----LRLAMDPPLTCNLNRGYAFVTF 210  
 Db 119 FIGGLRDLVGEEDLRLCCEIEIFVRATLIFVPHDILVKMDRSODSKGAYFAVF 178

Qy 211 CTKEAQAEAVLYNNHEIR-----SGKHIGVCIYANNRFLFGSISPKSKTKEQI 259  
 Db 179 KTKDVAQKATEELHSKEFKASSTANCSLSSLGKTIRCSLSETKONLFIGNIPKNWTEDBF 238

Qy 260 LBEFSKTVTEGTDVLYHQPDIDKKRNRFCPLEYBDHKTAQARERLMSGRVKWGNVGT 319  
 Db 239 RKVIEDVGPVGYENIELKDPTNTNNGFAFLVLYNNACADYSRKMDNSNFKLEGNAPT 298

Qy 320 VEADPFLDPEWMAKVYKLVFRNLANTYTEELEKAFSQGKLERVK-----LKD 372  
 Db 299 VTWADPKSPBSAAAQKVLYKVNIPENTSTEQKELFQRGKLERVK-----LKD 358

RESULT 4  
 DNZPRA  
 polyadenylate-binding protein - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 14-Feb-1992 #sequence\_revision 19-Jan-2001 #text\_change 19-Jan-2001  
 C;Accession: T38950; A39720  
 R;Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 Submitted to the EMBL Data Library, May 1997  
 A;Reference number: Z21819  
 A;Accession: T38950  
 A;Molecule type: DNA  
 A;Residues: 1-653 <SKE>  
 A;Cross-references: EMBL:Z99396; NID:96090525; PIDN:CA08762.1; PID:92104439; GSPDB:GN00

Qy 373 YAFIHEDRGAVKAMEEMNGKDLGENIEIVFAPPDKRKERKAQRQAKNQMYDDY 432  
 Db 359 FGFVHYAERSSALKAVKDTEREVNGQPLEVYLAKPQAERKHDPS-----Y 405

Qy 433 YYGPPHMPPPTRGRGRRGGYGPDDYGYEDYYDYGDYHNVRGGETDYYGYEDFQ 492  
 Db 406 SYGAAPTAPF-----VHPTFGFFAAAPYG-----A 431

Qy 493 VGARGRGRRGARGAAPSRGRGAAP-----PRGRAGY-SQRGGP-----529  
 Db 432 MGA-GUGAGTGSQPMYGRAMPMTQMVBNLPPGQGVYLVQGGMWAAAPPQRPRR 490

Qy 530 -----GSAR-----GVRAKGKRGR 542  
 Db 491 NDRNNGSSGGSGCRDNSHEDNGNRRYR 519

Result No.	Score	Query Match	Length	DB ID	Description
1	23.55	81.3	633	2 T02673	- heterogeneous nucleic acid binding protein
2	49.5	17.1	521	2 T01563	- hypothetical protein
3	43.5	15.0	471	2 T49019	- putative protein
4	287.5	9.9	653	1 DNZPPA	- putative protein
5	28.6	9.9	651	2 S18874	- putative protein
6	27.9	9.7	693	2 JC7925	- putative protein
7	27.9	9.6	671	2 C96534	- putative protein
8	27.8	9.6	705	2 S32644	- putative protein
9	27.5	9.5	345	1 B41732	- putative protein
10	27.4	9.5	651	2 T06979	- putative protein
11	26.8	9.3	623	2 T07933	- putative protein
12	26.6	9.2	662	2 T00497	- putative protein
13	25.8	8.9	629	2 T05425	- putative protein
14	25.8	8.9	713	2 A27441	- putative protein
15	25	8.7	405	2 H86249	- putative protein
16	251.5	8.7	638	2 S37085	- putative protein
17	251	8.7	448	2 T15542	- putative protein
18	24.9	8.6	353	1 S56750	- putative protein
19	24.9	8.6	692	2 T21095	- putative protein
20	24.7	8.5	500	2 S55785	- putative protein
21	24.4	8.4	707	2 A35804	- putative protein
22	242.5	8.4	636	2 I48708	- putative protein
23	24.2	8.4	427	2 T04823	- putative protein
24	24.1	8.3	712	2 JH0148	- putative protein
25	23.9	8.2	320	1 S20261	- putative protein
26	23.6	8.2	668	2 B96740	- putative protein
27	23.6	8.1	577	1 DNBYPA	- putative protein
28	23.4	8.1	574	2 S30887	- putative protein
29	23.3	8.0	308	1 DDRT	- putative protein

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6	30	233	8.0	320	1 A44485
OM protein - protein search, using sw model		31	233	8.0	320	2 S04617
Run on:	January 28, 2004, 09:23:10 ; Search time 20 Seconds (without alignments)	32	233	8.0	707	1 DNMS
Title:	US-09-978-242-1	33	232.5	8.0	414	2 JN0866
Perfect score:	2898	34	232.5	8.0	633	1 DNXLPA
Sequence:	1 MATEHVNGNGTTEPMDTTSA.....RGPGSARGVAGKGRGRGRS 545	35	231	8.0	566	2 T21096
Scoring table:	BLOSUM62	36	229	7.9	694	1 DNCHNL
Gapop:	10.0 , Gapext 0.5	37	227.5	7.9	628	2 S44138
Searched:	283308 seqs, 96168682 residues	38	227	7.8	522	2 S52431
Total number of hits satisfying chosen parameters:	283308	39	226.5	7.8	633	1 DNHUBA
Minimum DB seq length: 0	40	225	7.8	301	2 JW0079	
Maximum DB seq length: 2000000000	41	223	7.7	320	2 S30192	
Post-processing: Minimum Match 0%	42	222.5	7.7	646	2 T26427	
Maximum Match 100%	43	222	7.6	655	2 T07768	
Listing First 45 summaries	44	221	7.6	609	2 B84783	
Database :	PIR 76:*	45	218.5	7.5	308	2 B47369
	1: Pirl1:*					
	2: pir2:*					
	3: pir3:*					
	4: pir4:*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES	\$	Query	Match	Length	DB ID	Description
RESULT 1		T02673	heterogeneous nuclear ribonucleoprotein R - human			
C.Species: Homo sapiens (man)						
C;Accession: T02673						
R;Chan, E.K.L.; Mathison, D.A.; Portman, D.; Dreyfuss, G.; Steiner, G.; Tan, E.M.; Hassett, Nucleic Acids Res. 26, 439-445, 1998						
A;Title: Molecular definition of heterogeneous nuclear ribonucleoprotein R (hnRNP R) using a reference number: Z14697; PMID: 98083170; PMID: 9421497						
A;Accession: T02673						
A;Status: preliminary; translated from GB/EMBL/DDJB						
A;Molecule type: mRNA						
A;Residues: 1-633 <CHA>						
A;Cross-references: EMBL:AF00364; PID: AAC39540.1; PID: 92697102; PID: 92697103						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						
Matches 449; Conservative 51; Mismatches 40; Indels 20; Gaps 10;						
QY 1 MATRHVNNGNT---EFPMDTTSAVIHSENFTQLDAGLPQVAEKLDEIYVAGLYVHS 55						
Db 1 MANQ-VNGNAVQLKEEEFPMD-TSSVTHTEHYKTLIEAGLPQVAEKLDEIFQTGLVAYV 58						
Query Match 81.3%; Score 2355; DB 2; Length 633;						
Best Local Similarity 80.2%; Pred. No. 5.1e-143;						

Query Match		NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).	
Best Local Similarity	80.2%	RNA-BINDING (RRM) 1.	
Matches 449;	Conservative	RNA-BINDING (RRM) 2.	
		RNA-BINDING (RRM) 3.	
		RNA-BINDING (RGG-BOX).	
		X 11 AA APPROXIMATE REPEATS OF D-D-Y-Y.	
		G-Y-D-Y-H-D-Y.	
REPEAT	462	471	(APPROXIMATE).
REPEAT	472	482	2
REPEAT	488	497	3 (APPROXIMATE).
DOMAIN	579	633	GLN ASN RICH DOMAIN.
SEQUENCE	633 AA;	670943 MW;	089341F6465F0D46F CRC64;
Qy	81.3%;	Score: 2355;	DB 1; Length: 633;
Ddb	Pred. No.: 8.9e-138;	Mismatches: 51;	Indels: 20; Gaps
Qy	1 MATEHVNNGT - - - EEPMDTTSAVHSSENQFTLIDQCLPQYKAELDEIYTAVLYAHS :		
Ddb	1 MANQ - VNGHAVLQKKEEPMID - TSSVTHTEHYKTLIEAGLPQYKAELDEIYTQGTVAYV :		
Qy	56 DLDERIAEALKENEDGALAVLQQFKDSLJSHVNOKSAFLCGWMTKTYOREKOSTKVDAS :		
Ddb	59 DLDEDAIDALREFEEGALSVLQOOPKESDLJSHVNOKSAFLCGWMTKTYOREKOSKVDQS :		
Qy	116 SKGPDEAKIKALLERTGYTLDVITGQKRYGGPPDPDSVYSSGQOPSVGTETLVFGKIPRDLFE :		
Ddb	119 TKGPDDEAKIKALLERTGYTLDVITGQKRYGGPPDPDSVYSSGVPQSTIGEVFGKIPRDLFE :		
Qy	176 DELPLFEKAGPIPDLRLAMDPLTGLNRGTAFVIFCTCREAAQEAVKLNNHEIRSGRHLG :		
Ddb	179 DELPLFEKAGPIPDLRLAMDPLSQNRGYAFITFCGEAAQEAVKLCDSYEIRPGKHLG :		
Qy	236 VCISVANNRLFVGSIPKSKTKIEEFSKVTTEGLTDVLYHQDDKKCKNRGCFLEYED :		
Ddb	239 VCISVANNRLFVGSIPIKNTKRNLEBEFSKVTTEGLTDVLYHQDDKKCKNRGCFLEYED :		
Qy	296 HKTAQQAARRRLMSGVKVKVNGVTVIEWADPEVMAKVVLFVRNLANTVTEEL :		
Ddb	299 HKSAQAARRRLMSGVKVKVNGVTVIEWADPEVMAKVVLFVRNLANTVTEEL :		
Qy	356 KAFSGQFKLERVRQLKDYAIFIHDGRGAVKAMEMNGDLEGINIEIVFAKPDQKRKE :		
Ddb	359 KSFSEFGKLERVKLKDYAFVHFEDRGAAVKADMENGKEIEGEIEIVLAQPDQKRKE :		
Qy	416 RKAQROQAAKNOMYDDYYYYGPPHMBPTGRGR - GGRRGGYGYPPDYYGYEDYY :		
Ddb	419 RQAARQASRTAYEDYYHHPRMPEPIRGRRGGGRGGYPPDYYGYEDYYDYYGYD :		
Qy	474 YHNRYGGYESPPYYGYED - FOVGARORGGRGARGA - APSRGRAAAPPGRGAGSRGGP G :		
Ddb	479 YHDTRGGYEPPYYGYDDGAVARGRG - GGRCGRGAPPGRGAGPRGRGAGSRGAUPG :		
Qy	531 SARGTRAGK - - - - RGRG 543		
Ddb	538 PPRGSRGRRGGPAQQQRGRG 557		
RESULT 2			
PABP	SCHPO	STANDARD;	PRT;
ID	P31209 ; P87135 ;		653 AA.
AC	Created)		
DT	01-JUL-1993 (Rel. 26,		
DT	15-JUL-1998 (Rel. 36,		Last sequence update)
DT	28-FEB-2003 (Rel. 41,		Last annotation update)
DE	Polynucleotide-binding protein (Poly(A)-binding protein) (PABP).		
PABP	OR PABP OR SPAC57A-04C.		
GN	Schizosaccharomyces pombe (Fission yeast).		
OS	Eukaryota; Fungi; Ascomycota; Schizosaccharomyces; Schizosaccharomycetales; Schizosaccharomyces.		
OC	NCBI TaxID=4896 ;		
OC			

Result No.	Score	Query	Match	Length	DB ID	Description
1	2355	81.3	633	1	ROR_HUMAN	043290 homo sapien
2	2875	9.9	653	1	PABP_SCHPO	P31209 schizosaccharomyces pombe
3	286	9.9	650	1	NUCL_XENLA	P20397 xenopus laevis
4	274	9.5	345	1	SQD_DROME	Q08473 drosophila melanogaster
5	260.5	9.0	644	1	PAB4_HUMAN	Q13310 homo sapien
6	258.5	8.9	629	1	PAB2_ARATH	P4231 arabidopsis thaliana
7	258	8.9	713	1	NUCL_MESAU	P08199 mesocyclospora ulei
8	248	8.6	631	1	PAB3_HUMAN	Q91361 homo sapien
9	247	8.5	500	1	GAR2_SCHPO	P41891 schizosaccharomyces pombe
10	244	8.4	706	1	NUCL_HUMAN	P19338 homo sapien
11	242.5	8.4	636	1	PAB1_MOUSE	P23341 mus musculus
12	241.5	8.3	712	1	NUCL_RAT	P13383 rattus norvegicus
13	240.5	8.3	636	1	PAB1_HUMAN	P11940 homo sapien
14	236.5	8.2	668	1	PAB5_ARATH	Q05196 arabidopsis thaliana
15	236	8.1	576	1	PABP_YEAST	P04147 saccharomyces cerevisiae
16	234	8.1	632	1	PABP_DROME	P21187 drosophila melanogaster
17	233	8.0	319	1	ROA1_MOUSE	P49312 mus musculus
18	233	8.0	706	1	NUCL_MOUSE	P09405 mus musculus
19	232.5	8.0	414	1	NOP3_YEAST	Q01560 saccharomyces cerevisiae
20	232.5	8.0	633	1	PAB1_XENLA	P20965 xenopus laevis
21	231	8.0	319	1	ROA1_RAT	P04256 rattus norvegicus
22	229	7.9	391	1	ROG_HUMAN	P38159 homo sapien
23	229	7.9	694	1	NUCL_CHICK	P15771 gallus gallus
24	227	7.8	522	1	PAB2_HUMAN	Q15097 homo sapien
25	224.5	7.7	660	1	PAB3_ARATH	Q61380 arabidopsis thaliana
26	223	7.7	319	1	ROA1_MACMU	Q28521 macaca mulatta
27	222.5	7.7	353	1	ROD_FAT	Q93154 rattus norvegicus
28	222	7.7	371	1	ROA1_HUMAN	P03651 homo sapien
29	221.5	7.6	379	1	ROA3_MOUSE	Q88905 mus musculus
30	221	7.6	609	1	PABX_ARATH	Q92948 arabidopsis thaliana
31	219.5	7.6	305	1	ROA0_HUMAN	Q13151 homo sapien
32	219.5	7.6	355	1	ROD_HUMAN	Q14103 homo sapien
33	218.5	7.5	424	1	SD3B4_HUMAN	Q15427 homo sapien

Db	121 BAKIKALLERTGTYLDTQORKYGGPPPSVSGQQPSVGTEIFVGKIPRDLFEDELVP 180	Qy	61 AIEALKKEFNEGDALAVLQQFKOSDLHIVONKSAAFLCGVMKTYQRREKOGTKVADSKGPD 120
Qy	181 LFEKAGPINDLRLIMMDPLTCLSLNRYGAFVTFECTKEAAQEAVKLVNHEITRSKGHIGCVTSV 240	Db	61 AIEALKKEFNEGDALAVLQQFKOSDLHIVONKSAAFLCGVMKTYQRREKOGTKVADSKGPD 120
Db	181 LFEKAGPINDLRLIMMDPLTCLSLNRYGAFVTFECTKEAAQEAVKLVNHEITRSKGHIGCVTSV 240	Qy	121 BAKIKALLERTGTYLDTQORKYGGPPPSVSGQQPSVGTEIFVGKIPRDLFEDELVP 180
Qy	241 ANNRLFGS1PKSKTKEQ1LLEFSKVTGILTDLHYHQDDKKNGPFFLEYEDHKTA 300	Db	121 BAKIKALLERTGTYLDTQORKYGGPPPSVSGQQPSVGTEIFVGKIPRDLFEDELVP 180
Db	241 ANNRLFGS1PKSKTKEQ1LLEFSKVTGILTDLHYHQDDKKNGPFFLEYEDHKTA 300	Qy	181 LFEKAGPIWDLRLIMMDPLTGLNRGYAFTFCTKEAAQEAVLYNNHEIRSGKHIGCVTSV 240
Qy	301 QARRRLMSGRVKYKVGNGVGTGEWAIDPDEPEWAKVYLVRNLANTVTEELKEAFSQ 360	Db	181 LFEKAGPFWDLRLIMMDPLTGLNRGYAFTFCTKEAAQEAVLYNNHEIRSGKHIGCVTSV 240
Db	301 QARRRLMSGRVKYKVGNGVGTGEWAIDPDEPEWAKVYLVRNLANTVTEELKEAFSQ 360	Qy	241 ANNRLFVGS1PKSKTKEQ1LLEFSKVTGILTDLHYHQDDKKNGPFFLEYEDHKTA 300
Qy	361 FGKLERVKLKDKYDIAFHFDERGAVKAMEEMNGKDLLEGENEIYFAKPPDKRKERKACR 420	Db	241 ANNRLFVGS1PKSKTKEQ1LLEFSKVTGILTDLHYHQDDKKNGPFFLEYEDHKTA 300
Db	361 FGKLERVKLKDKYDIAFHFDERGAVKAMEEMNGKDLLEGENEIYFAKPPDKRKERKACR 420	Qy	301 QARRRLMSGRVKYKVGNGVGTGEWAIDPDEPEWAKVYLVRNLANTVTEELKEAFSQ 360
Qy	421 QAAKNQMYDDYYGGPPMPPTGRGRGRRGGYGYPPDYYGYDYDHNYRG 480	Db	301 QARRRLMSGRVKYKVGNGVGTGEWAIDPDEPEWAKVYLVRNLANTVTEELKEAFSQ 360
Db	421 QAAKNQMYDDYYGGPPMPPTGRGRGRRGGYGYPPDYYGYDYDHNYRG 480	Qy	361 FGKLERVKLKDKYDIAFHFDERGAVKAMEEMNGKDLLEGENEIYFAKPPDKRKERKACR 420
Qy	481 YEDPYGYEDEQVQARGRGERGARGAAAPSQRGRGAAPPRGRAGYSQRGGSARGVRAGK- 539	Db	361 FGKLERVKLKDKYDIAFHFDERGAVKAMEEMNGKDLLEGENEIYFAKPPDKRKERKACR 420
Db	481 YEDPYGYEDEQVQARGRGERGARGAAAPSQRGRGAAPPRGRAGYSQRGGSARGVRAGK 540	Qy	421 QAARKNQMYDDYYGGPPMPPTGRGRGRRGGYGYPPDYYGYDYDHNYRG 480
Qy	540 ----RGRGR 544	Db	421 QAARKNQMYDDYYGGPPMPPTGRGRGRRGGYGYPPDYYGYDYDHNYRG 480
Db	541 GAQQQRGRGQ 550	Qy	481 YEDPYGYEDEQVQARGRGERGARGAAAPSQRGRGAAPPRGRAGYSQRGGSARGVRAGK- 539
Qy	540 ----RGRGR 543	Db	481 YEDPYGYEDEQVQARGRGERGARGAAAPSQRGRGAAPPRGRAGYSQRGGSARGVRAGK 540
RESULT 2			
Q9QYF4	PRELIMINARY; PRT; 561 AA.	Q9QYF4	PRELIMINARY; PRT; 561 AA.
ID	Q9QYF4	AC	Q9QYF4;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)	DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE	SYNCRIP protein.	DE	SYNCRIP protein.
GN	NSAPI OR SYNCRIPI.	GN	NSAPI OR SYNCRIPI.
OS	Mus musculus (Mouse)	OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Muriinae; Mus. NCBI_TaxID=10900; [1]
OC		OC	
OX		OX	
RN	SEQUENCE FROM N.A.	RN	SEQUENCE FROM N.A.
RP	STRAIN=DDV;	RP	Homo sapiens (Human).
RX	MEDLINE=20200483; PubMed=10734137;	RX	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RA	Mizutani A., Fukuda M., Ibara K., Shiraiishi Y., Mikoshiba K.;	RA	Harris C.E., Boden R.A., Arell C.R.;
RT	"SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptoctagmin isoforms."	RT	"A novel heterogeneous nuclear ribonucleoprotein-like protein interacts with NS1 of the minute virus of mice."
RT	InterPro: IPR000504; RNA_recC_mot.	RT	interacts with NS1 of the minute virus of mice."
DR	SMART: SM00360; RNA_recC.	DR	[1] J. Virol. 73:72-80 (1999).
DR	InterPro: IPR006535; hnRNP_R_Q.	DR	EMBL: AP155568; AACD38198.1; -.
DR	MGD: 14891620; RRM; 3.	DR	HSSP: P19339; 2SXL.
DR	InterPro: IPR000504; RNA_recC_mot.	DR	InterPro: IPK000504; hnRNP_R_Q.
DR	SMART: SM00360; RNA_recC.	DR	InterPro: IPK000504; RNA_recC_mot.
DR	TIGRFAMS: TIGR01648; hnRNP_R-Q; 1.	DR	Pfam: PF00076; rrm; 3.
DR	PROSITE: PS00030; RRM; 3.	DR	SMART: SM00360; RRM; 3.
DR	PROSITE: PS00030; RRM_RNP_1; 2.	DR	TIGRFAMS: TIGR01648; hnRNP_R-Q; 1.
SQ	SEQUENCE 561 AA; 62514 MW; C0259CS4014616A CRC64;	SQ	DR PROSITE: PS50102; RRM; 3.
Query Match	98.7%; Score 2860; DB 11; Length 561;	DR PROSITE: PS00030; RRM_RNP_1; 2.	DR PROSITE: PS50102; RRM; 3.
Best Local Similarity	98.2%; Pred. No. 1..7e-191; Indels 6; Gaps 1;	SEQUENCE 562 AA; 62656 MW;	SEQUENCE 562 AA; 62656 MW;
Matches	Conservative 2; Mismatches 2;	Query Match 98.7%; Best Local Similarity 98.0%; Pred. No. 1..7e-191; Mismatches 3; Conservative 3; Indels 6; Gaps 1;	Best Local Similarity 98.0%; Pred. No. 1..7e-191; Mismatches 3; Conservative 3; Indels 6; Gaps 1;
Qy	1 MATEHYNNGTEEPMDTTSAVHSENFTQLDAGLPQKVAEKLDEIYVAGLVLAISDLDR 60	Qy	1 MATEHYNNGTEEPMDTTSAVHSENFTQLDAGLPQKVAEKLDEIYVAGLVLAISDLDR 60
Db	1 MATEHYNNGTEEPMDTTSAVHSENFTQLDAGLPQKVAEKLDEIYVAGLVLAISDLDR 60	Db	

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2861	98.7	562	11	Q8BGP1	Q8bgp1	mus musculu
2	2860	98.7	561	11	Q9QFP4	Q9qfp4	mus musculu
3	2860	98.7	562	4	Q9YSB3	Q9y5b3	homo sapien
4	2860	98.7	627	11	Q8CCC2	Q8ccg2	mus musculu
5	2857	98.6	561	4	Q96LC1	Q96lc1	homo sapien
6	2857	98.6	623	4	060506	060506	homo sapien
7	2823	97.6	625	11	Q8BS91	Q8bs91	mus musculu
8	2647	91.4	527	4	Q9Y599	Q9y599	homo sapien
9	2640	91.4	588	4	Q96LC2	Q96lc2	homo sapien
10	2533	87.4	558	11	Q91ZRO	Q91zro	mus musculu
11	2447	85.8	491	11	Q8VEM6	Q8cem6	mus musculu
12	2365	81.6	632	11	Q8VEM5	Q8cem5	mus musculu
13	2343	80.9	636	4	Q9BV64	Q9bv64	homo sapien
14	2227	76.8	601	11	Q99KG1	Q99kg1	mus musculu
15	2043	69.8	410	4	Q8IW78	Q8iw78	homo sapien
16	1667	57.5	380	11	Q9CT37	Q9ct37	mus musculu

## ALIGNMENTS

RESULT	1	QBGP1	QBGP1;	PRELIMINARY;	PRT;	562 AA.
		ID	QBGP1;			
		AC	QBGP1;			
		DT	01-MAR-2003	(TrEMBLrel. 23, Created)		
		DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)		
		DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
		DE				
		OS				
		OC				
		RN	[1]			
		RP				
		RC	SEQUENCE FROM N.A.			
		RC	STRAIN=C57BL/6J; TISSUE=Body, and Embryo;			
		RX	MEDLINE=22334683; PubMed=12466651;			
		RA	The FANTOM Consortium,			
		RA	the RIKEN Genome Exploration Research Group Phase I & II Team;			
		RT	"Analysis of the mouse transcriptome based on functional annotation of			
		RT	60,770 full-length cDNAs."			
		RL	Nature 420:563-573 (2002)			
		DR	EMBL; AK034845; BAC28852; EMBL; AK077588; BAC36880; -			
		DR	SEQUENCE 562 AA; 62672 MW; 8103CFA286105377 CRC64;			
		DR	98.7% Score 2861; DB 11; Length 562;			
		DR	1 MATEHYNQNGTEPMDITSAVHSNFQTLIDAGLPQKVAEKLDETVAGLYAHSDILDE 60			
		DR	1 MATEHYNQNGTEPMDITSAVHSNFQTLIDAGLPQKVAEKLDETVAGLYAHSDILDE 60			
		QY	61 AIEALKFNEQDALAVLQQPKQDSLIVQNSKAFLQGYMCKTYRQEKGTYADSKGPD 120			
		Db	1 MATEHYNQNGTEPMDITSAVHSNFQTLIDAGLPQKVAEKLDETVAGLYAHSDILDE 60			
		QY	61 AIEALKFNEQDALAVLQQPKQDSLIVQNSKAFLQGYMCKTYRQEKGTYADSKGPD 120			
		Db	1 MATEHYNQNGTEPMDITSAVHSNFQTLIDAGLPQKVAEKLDETVAGLYAHSDILDE 60			
		QY	121 EAKIKALLERTGTYLTDYTGTGKGGPPDVSYSGQQPSVGTEIFVKIIPDLFEDLV P 180			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2861	98.7	562	11	Q8BGP1	Q8bgp1	mus musculu
2	2860	98.7	561	11	Q9QFP4	Q9qfp4	mus musculu
3	2860	98.7	562	4	Q9YSB3	Q9y5b3	homo sapien
4	2860	98.7	627	11	Q8CCC2	Q8ccg2	mus musculu
5	2857	98.6	561	4	Q96LC1	Q96lc1	homo sapien
6	2857	98.6	623	4	060506	060506	homo sapien
7	2823	97.6	625	11	Q8BS91	Q8bs91	mus musculu
8	2647	91.4	527	4	Q9Y599	Q9y599	homo sapien
9	2640	91.4	588	4	Q96LC2	Q96lc2	homo sapien
10	2533	87.4	558	11	Q91ZRO	Q91zro	mus musculu
11	2447	85.8	491	11	Q8VEM6	Q8cem6	mus musculu
12	2365	81.6	632	11	Q8VEM5	Q8cem5	mus musculu
13	2343	80.9	636	4	Q9BV64	Q9bv64	homo sapien
14	2227	76.8	601	11	Q99KG1	Q99kg1	mus musculu
15	2043	69.8	410	4	Q8IW78	Q8iw78	homo sapien
16	1667	57.5	380	11	Q9CT37	Q9ct37	mus musculu

source	1. .2079	/organism="unknown"	535	9	528	t
BASE COUNT	619	a	377	c	528	t
ORIGIN						
Query Match	100.0%	Score 2079; DB 6; Length 2079;				
Best Local Similarity	100.0%	Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;				
Matches 2079; Conservative						
Qy	1	GGGGCGCGGGCACCGGAGCCGGCTGGAGCGACTGGATCGGGTTCGGCT 60				
Db	1	GGGGCGCGGGCACCGGAGCCGGCTGGAGCGACTGGATCGGGTTCGGCT 60				
Qy	61	CCAGGGCGTGAGCTTGCGGGCATTTACACAGCTCCACTGGGGCACAGGA 120				
Db	61	CCAGGGCGTGAGCTTGCGGGCATTTACACAGCTCCACTGGGGCACAGGA 120				
Qy	121	GCAGGGAGGACGGACGGCTTCCGGCAACCCGATAACATCGGACAGGATTCTCGGCC 180				
Db	121	GCAGGGAGGACGGACGGCTTCCGGCAACCCGATAACATCGGACAGGATTCTCGGCC 180				
Qy	181	CAACGGGGAGATCTGGAAACATGGCTACAGAACATGTAAATGGTACTGAGA 240				
Db	181	CAACGGGGAGATCTGGAAACATGGCTACAGAACATGTAAATGGTACTGAGA 240				
Qy	241	GCCATGGAPACTACTTCGAGTTACCATTCGAGTTACAGCTTTGCTGTATGC 300				
Db	241	GCCATGGAPACTACTTCGAGTTACCATTCGAGTTACAGCTTTGCTGTATGC 300				
Qy	301	TGGTTACACAGAAAGTTCTGAAAAACTATGATGAAATTACGTTGAGGGTAGITTC 360				
Db	301	TGGTTACACAGAAAGTTCTGAAAAACTATGATGAAATTACGTTGAGGGTAGITTC 360				
Qy	361	ACATAGTGATTATGATGAAAGAGCTTAAAGAACATGGCTTAAGAACAGGGTTC 420				
Db	361	ACATAGTGATTATGATGAAAGAGCTTAAAGAACATGGCTTAAGAACAGGGTTC 420				
Qy	421	ATGGCAGTTCTCACAGTTAACAGCTTAAGACAGCTCATCTCATGTTAACAAAGTC 480				
Db	421	ATGGCAGTTCTCACAGTTAACAGCTTAAGACAGCTCATCTCATGTTAACAAAGTC 480				
Qy	481	CTTTTTATGTTGGAGCTCATGAGACTTACAGGAGAAAACAGGGGACAAAGTAC 540				
Db	481	CTTTTTATGTTGGAGCTCATGAGACTTACAGGAGAAAACAGGGGACAAAGTAC 540				
Qy	541	AGATCTTGTAAAGGACCAGTGTGGAAAGAACAGGGCAAAAGCTTCACTGGT 600				
Db	541	AGATCTTGTAAAGGACCAGTGTGGAAAGAACAGGGCAAAAGCTTCACTGGT 600				
Qy	601	CACACTTGAGTGACCACTGGACACTGGACAGGGAGTATGGGACCACTCCGTTA 660				
Db	601	CACACTTGAGTGACCACTGGACACTGGACAGGGAGTATGGGACCACTCCGTTA 660				
Qy	661	TTCAAGTCAGTCAGCTTCTGTGGACTGGAGATTTGTGGAAAGATCCAAAGAGCT 720				
Db	661	TTCAAGTCAGTCAGCTTCTGTGGACTGGAGATTTGTGGAAAGATCCAAAGAGCT 720				
Qy	721	ATTTGGGGATGCACTGCACTGCAACTGTGTCATTATGGAAAAGCTGGCTTCTG 780				
Db	721	ATTTGGGGATGCACTGCACTGCAACTGTGTCATTATGGAAAAGCTGGCTTCTG 780				
Qy	781	AATGATGGATCCACTCACTGGTCAATAGAGCTTATGGCTTGTCACTTTTGTCACA 840				
Db	781	AATGATGGATCCACTCACTGGTCAATAGAGCTTATGGCTTGTCACTTTTGTCACA 840				
Qy	841	AGAAGCAGCTCAGGGCTGGTTAACTGTATAAATATGAAATTGTCTGAAAACA 900				
Db	841	AGAAGCAGCTCAGGGCTGGTTAACTGTATAAATATGAAATTGTCTGAAAACA 900				
Qy	901	TATTGGTGCTGCACTGCACTGCAACTAGGCTTCTGCAACAACTGGCTTCTG 960				
Db	901	TATTGGTGCTGCACTGCACTGCAACTAGGCTTCTGCAACAACTGGCTTCTG 960				
Qy	961	TAAACCAGGAAACAGATTCTGAAAGAATTAGCAAAAGTAACAGGGCTTACAGACGT 1020				
Db	961	TAAACCAGGAAACAGATTCTGAAAGAATTAGCAAAAGTAACAGGGCTTACAGACGT 1020				
Qy	1021	CATTATACACCAACCGATGACAGAAACAAACAGGGCTTTCGCTTCTGATA 1080				
Db	1021	CATTATACACCAACCGATGACAGAAACAAACAGGGCTTTCGCTTCTGATA 1080				
Qy	1081	TGAGATCACAAAACAGTGCCTGGAGGGCTAGGTAAATGAGTGTAAAGTCAGGT 1140				
Db	1081	TGAGATCACAAAACAGTGCCTGGAGGGCTAGGTAAATGAGTGTAAAGTCAGGT 1140				
Qy	1141	CTGGGGAAATGTGGBACTGTGAACTGGCTGATCTATAGAGATCCTGATCTG 1200				
Db	1141	CTGGGGAAATGTGGBACTGTGAACTGGCTGATCTATAGAGATCCTGATCTG 1200				
Qy	1201	TATGCCAAGTTAAAGTAAAGTGTGCTGTTAGCTGAACTTGTCAACAGAAAGAT 1260				
Db	1201	TATGCCAAGTTAAAGTGTGCTGTTAGCTGAACTTGTCAACAGAAAGAT 1260				
Qy	1261	TTTAGAAAGGCAATTACTCGTTGAGCTTACTGAGTAAAGGTTAAAGGATTA 1320				
Db	1261	TTTAGAAAGGCAATTACTCGTTGAGCTTACTGAGTAAAGGTTAAAGGATTA 1320				
Qy	1321	TGCGTCATTCATTGATGAGGGAGATGGCTTACGGCTATGGAGAAATGATG 1380				
Db	1321	TGCGTCATTCATTGATGAGGGAGATGGCTTACGGCTATGGAGAAATGATG 1380				
Qy	1381	CAAGACTTGGAGGGAGAAATTGAGATTGTTGCTGCAAGCCACCGATCAGAAAG 1440				
Db	1381	CAAGACTTGGAGGGAGAAATTGAGATTGTTGCTGCAAGCCACCGATCAGAAAG 1440				
Qy	1441	GAAGAAAGAAAAGCTCAGAGGAAGCAGCAAAATAAATGATGAGTAACTACTA 1500				
Db	1441	GAAGAAAGAAAAGCTCAGAGGAAGCAGCAAAATAAATGATGAGTAACTACTA 1500				
Qy	1501	TTATGTCACCTCATATGCCCTCCACAAAGGTTGGGGCTGGAGGTAGGTG 1560				
Db	1501	TTATGTCACCTCATATGCCCTCCACAAAGGTTGGGGCTGGAGGTAGGTG 1560				
Qy	1621	TTACCCATAACTATGTGGTATGAGATCCATACTATGGTTGAGATTTCACTG 1680				
Db	1621	TTACCCATAACTATGTGGTATGAGATCCATACTATGGTTGAGATTTCACTG 1680				
Qy	1681	TGGAGCTAGAGGAAGGGGGTTAGAGGGCAAGGGCTGGCTCATCCAGGTG 1740				
Db	1681	TGGAGCTAGAGGAAGGGGGTTAGAGGGCAAGGGCTGGCTCATCCAGGTG 1740				
Qy	1801	AGGGCTTCAGAGGGAAAGGGGTGAGGCCGTTCTGAGGCCGTTCTGAGACTCA 1860				
Db	1801	AGGGCTTCAGAGGGAAAGGGGTGAGGCCGTTCTGAGGCCGTTCTGAGACTCA 1860				
Qy	1861	CTTGCTCATGTGGATTAACAGGAGCTTCCTGAGCTTCTGAGCTTCTGAGTC 1920				
Db	1861	CTTGCTCATGTGGATTAACAGGAGCTTCCTGAGCTTCTGAGCTTCTGAGTC 1920				
Qy	1921	CCTTAATAATGTCGGCTGATAGGAGCATATTCTGAGAAAGACCTTCCTGAGTC 1980				
Db	1921	CCTTAATAATGTCGGCTGATAGGAGCATATTCTGAGAAAGACCTTCCTGAGTC 1980				
Qy	1981	CATGGAAATAAACATGGGCAATTGAGCTTACATGGAACTTGTGAGTTAACCTACT 2040				
Db	1981	CATGGAAATAAACATGGGCAATTGAGCTTACATGGAACTTGTGAGTTAACCTACT 2040				
Qy	2041	AATTTCCTCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 2079				

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description	
1	2079	100.0	2079	6	AR177851		ARI177851 Sequence	
2	2012.6	96.8	2221	9	AF155568	Homo sapi	AF155568 Sequence	
3	1883.2	90.6	2208	10	BC041148	Mus muscu	BC041148 Sequence	
4	1848.4	88.9	103819	9	HSU775C13	Human DNA	AL109618 Sequence	
5	1848.4	88.9	212659	2	AL390737	Homo sapi	AL390737 Sequence	
6	1826.2	87.8	3452	6	BD143848	RNA-bind	BD143848 Sequence	
7	1826.2	87.8	3452	10	AB035725	Mus muscu	AB035725 Sequence	
8	1810.6	87.1	2932	9	AF037448	Homo sapi	AF037448 Sequence	
9	1809.2	87.0	3671	10	BC050079	Mus muscu	BC050079 Sequence	
10	1695.4	81.5	3319	9	BC032643	Homo sapi	BC032643 Sequence	
11	1619.8	77.9	1941	9	BC040844	Homo sapi	BC040844 Sequence	
12	1610.2	77.5	1686	9	AY034483	Homo sapi	AY034483 Sequence	
13	1610.2	77.5	1872	9	AY034481	Homo sapi	AY034481 Sequence	
14	1502.8	72.3	2373	10	AF093821	Mus muscu	AF093821 Sequence	
15	1463.4	70.4	3101	6	AK011753	Sequence	AK011753 Sequence	
16	1390.2	66.9	1767	9	AY034482	Homo sapi	AY034482 Sequence	
17	1329.2	63.9	2232	10	AF040843	Mus muscu	AF040843 Sequence	
18	1089.4	52.4	259969	2	AC118904	Rattus no	AC118904 Sequence	
c	19	1089.4	52.4	288143	2	AC109106	Rattus no	AC109106 Sequence
c	20	935.2	45.0	2371	5	BC046902	Danio rer	BC046902 Sequence
c	21	901.8	43.4	226602	2	AC122217	Mus muscu	AC122217 Sequence
c	22	895	43.0	1899	10	AF411128	Sequence	AF411128 Sequence
c	23	895	43.0	2725	10	BC038051	Mus muscu	BC038051 Sequence
c	24	884.4	42.5	1899	10	AY184814	Rattus no	AY184814 Sequence
c	25	882	42.4	2644	9	AF000364	Homo sapi	AF000364 Sequence
c	26	864.6	41.6	2722	9	BC001449	Human DNA	BC001449 Sequence
c	27	852.8	41.0	202374	10	BC046779	Mus muscu	BC046779 Sequence
c	28	846.4	40.7	235938	2	AC126148	Rattus no	AC126148 Sequence
c	29	846.4	40.7	264178	2	AC098650	Rattus no	AC098650 Sequence
c	30	792.4	38.1	934	6	AC103715	Sequence	AC103715 Sequence
c	31	763	36.7	196188	2	AC084410	Mus muscu	AC084410 Sequence
c	32	763	36.7	205520	2	AC079134	Mus muscu	AC079134 Sequence
c	33	705.6	33.9	202374	10	AC122868	Mus muscu	AC122868 Sequence
c	34	693.8	33.4	204962	2	AC127932	Rattus no	AC127932 Sequence
c	35	693.8	33.4	225370	2	AC129693	Rattus no	AC129693 Sequence
c	36	686.6	33.0	215801	2	AC140330	Mus muscu	AC140330 Sequence
c	37	642.2	30.9	139936	9	AL161799	Human DNA	AL161799 Sequence
c	38	482.8	23.2	2195	10	BC004001	Mus muscu	BC004001 Sequence
c	39	445.2	21.4	1236	3	AK116214	Ciona int	AK116214 Sequence
c	40	435.2	20.9	263129	2	AC098289	Human DNA	AC098289 Sequence
c	41	424.6	20.4	121127	2	AC119124	Rattus no	AC119124 Sequence
c	42	424.6	20.4	241168	2	AC105884	Rattus no	AC105884 Sequence
c	43	407.2	19.6	265259	2	AC121468	Rattus no	AC121468 Sequence
c	44	395.8	19.0	483	6	BD143847	RNA-bind	BD143847 Sequence
c	45	378.6	18.2	2606	3	AY058477	Drosophil	AY058477 Sequence

## ALIGNMENTS

RESULT	1	ARI177851	2	from patent	2079 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	ARI177851	Sequence	2	US 6313266.				
DEFINITION								
ACCESSION		ARI177851						
VERSION								
KEYWORDS								
SOURCE								
ORGANISM								
UNCLASSIFIED								
REFERENCE	1	(bases 1 to 2079)						
AUTHORS	Bandman, O.							
TITLE	Human nucleolin-like protein							
JOURNAL	Corley, N.C. and Shah, P.							
FEATURES	Location/Qualifiers							

Pred. No. is the number of results predicted by chance to have a a

Copyright (c) 1993 - 2004 Compugen Ltd.

Copyright (c) 1993 - 2004 Compugen Ltd.

The present sequence encodes human nucleolin-like peptide, designated HNLPL. HNLPL is the main protein component in the nucleolus of eukaryotic cells and is an essential part of ribosome biosynthesis and also plays an important role in importing proteins to the nucleus. HNLPL may be used in the diagnosis, prevention and treatment of disorders associated with abnormal expression of HNLPL. For example, it may be used to treat cancers (e.g. melanoma, breast cancer and prostate cancer), autoimmune diseases (e.g. autoimmune haemolytic anaemia and inflammatory bowel disease) and, in particular, Alzheimer's disease. It may also be used to study the function of the HNLPL peptides, the formation (biosynthesis) of the complex of ribosomes and nucleolus, the assembly of ribosomes and the removal of the nucleolus.

X	Q	Sequence 2079 BP; 639 A; 377 C; 535 G; 528 T; 0 other;	Qy	
	Query Match	100.0% ; Score 2079; DB 20; Length 2079;	Db	
	Best Local Similarity	100.0% ; Pred. No. 0;	Db	1021 CATTTTATACCAACCCGATACAGAAAACAGGGCTTTCCTTGATAA 108
	Matches 2079; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	Db	1021 CATTTTATACCAACCCGATACAGAAAACAGGGCTTTCCTTGATAA 108
b	1	GGGGCGCCGCCGCCGCCACCGGAGGCCCTCGGAGGGAGTGAACTGGATGGGTTTGCTG 60	Qy	1081 TGAGATCACAAACAGCTGCCAGGCTAGGCTTAAGCTTAAGTCAAAGT 114
b	1	GGGGCGCCGCCGCCACGGAGGCCCTCGGAGGGAGTGAACTGGATGGGTTTGCTG 60	Db	1081 TGAGATCACAAACAGCTGCCAGGCTAGGCTTAAGTCAAAGTCAAAGT 114
b	61	CCAGGGCGGTGAGCTGGGGGCATTAAACAGCTCCACTCGCCGAGACAGGGA 120	Qy	1141 CTGGGGAAATGTTGAACTGTTGAATGGCTGATCCTATAGAAGATCTTGATCTTGAGCT 120
b	61	CCAGGGCGGTGAGCTGGGGGCATTAAACAGCTCCACTCGCCGAGACAGGGA 120	Db	1141 CTGGGGAAATGTTGAACTGTTGAATGGCTGATCCTATAGAAGATCTTGATCTTGAGCT 120
b	121	GCAGGGCGACGGCGTCCCGCAACCGATACCATGACAGGATTCTCGCCCTCAGCC 180	Qy	1201 TATGCCAAGGTAAGTAAAGTGCTGTTGAGCAGGTTAACCTGTAACAGAGAT 126
b	121	GCAGGGCGACGGCGTCCCGCAACCGATACCATGACAGGATTCTCGCCCTCAGCC 180	Db	1201 TATGCCAAGGTAAGTAAAGTGCTGTTGAGCAGGATTCAACAGAGAT 126
b	181	CAACGGGAGATCTCGGAACATGGCTACAGAACATGTTAATGGTACTGAGA 240	Qy	1261 TTGAAAGGGCATTTAGTCAGTTGGAAACTCTGTTGAGCTTAAAGATTA 132
b	181	CAACGGGAGATCTCGGAACATGGCTACAGAACATGTTAATGGTACTGAGA 240	Db	1261 TTGAAAGGGCATTTAGTCAGTTGGAAACTCTGTTGAGCTTAAAGATTA 132
b	241	GCCCCATGATACTACTCTCGATGTTATCCATTAGAACATTTCAGAACATGCTGATGC 300	Qy	1321 TGGCTTCATTCATTTGATGAGCCAGATGGCTGCTCAAGGCTATGGAAATGAAATG 138
b	241	GCCCCATGATACTACTCTCGATGTTATCCATTAGAACATTTCAGAACATGCTGATGC 300	Db	1321 TGGCTTCATTCATTTGATGAGCCAGATGGCTGCTCAAGGCTATGGAAATGAAATG 138
b	301	TGGTTAACACAGAACATTAGTGAATAATTAGTGAATAATTTCAGAACATTTCAGAACATGCTGATGC 360	Qy	1381 CAAGACATTTGGAGGAGAAAATTTGAAATTGTTTTGCCAACCCAGATGCAAAG 144
b	301	TGGTTAACACAGAACATTAGTGAATAATTAGTGAATAATTTCAGAACATTTCAGAACATGCTGATGC 360	Db	1381 CAAGACATTTGGAGGAGAAAATTTGAAATTGTTTTGCCAACCCAGATGCAAAG 144
b	361	ACATAGTGTAGTAAAGGCTATTGAGCTTAAAGAATTCAATGAAAGGGTGC 420	Qy	1441 GAAAGAAAAGCTCAGGGCAAGCAAAATCAATGTTGAGCTTACTA 150
b	361	ACATAGTGTAGTAAAGGCTATTGAGCTTAAAGAATTCAATGAAAGGGTGC 420	Db	1441 GAAAGAAAAGCTCAGGGCAAGCAAAATCAATGTTGAGCTTACTA 150
b	421	ATTGGCGATTCTTCACAGTTAAAGACAGTCACTCTCATGTTCAAGAACAAAGTGC 480	Qy	1501 TTATGGTCCACCTCATATGCCCTTCACAAGGGCTGAGGGTAGGGCTGG 156
b	421	ATTGGCGATTCTTCACAGTTAAAGACAGTCACTCTCATGTTCAAGAACAAAGTGC 480	Db	1501 TTATGGTCCACCTCATATGCCCTTCACAAGGGCTGAGGGTAGGGCTGG 156
b	481	CCTTTTATGTGGAGTCATGAGAACATCAAGGGAGAGAAAACAGGGACCAAGTGC 540	Qy	1561 TTATGGTATTCCTCCAGATTATTTGATATGAGATTATTATGTTTATCA 162
b	481	CCTTTTATGTGGAGTCATGAGAACATCAAGGGAGAGAAAACAGGGACCAAGTGC 540	Db	1561 TTATGGTATTCCTCCAGATTATTTGATATGAGATTATTATGTTTATCA 162
b	541	AGATTCTAGTAAAGGACCCAGTGGGAAATTAGGCACTCTGGAAACAGGGCTA 600	Qy	1621 TTACCAATAACTATCTGGGATATGAGATTCTGTTATGAGATTTCAGT 168
b	541	AGATTCTAGTAAAGGACCCAGTGGGAAATTAGGCACTCTGGAAACAGGGCTA 600	Db	1621 TTACCAATAACTATCTGGGATATGAGATTCTGTTATGAGATTTCAGT 168
b	601	CAACATGTGTGACCACTGGACAGACGGAAAGTATGGAGCCACCTCTCAGATTCCGTTA 660	Qy	1681 TGGAGCTAGGGAAAGGGCTGGTAGGGACCAAGGGCTGTCATCCAGAGTCTGTC 174
b	601	CAACATGTGTGACCACTGGACAGACGGAAAGTATGGAGCCACCTCTCAGATTCCGTTA 660	Db	1681 TGGAGCTAGGGAAAGGGCTGGTAGGGACCAAGGGCTGTCATCCAGAGTCTGTC 174
b	1741	GGCTGGTCTCCCGCCGGTAGAGGGGGTATTCAAGAGGGAGCTGGATAGGCAAG 184	Qy	

CDNA downregulated 25 ACA03964  
 CDNA sequence #199 ABX35808  
 Human breast cancer 24 AAL2671  
 Human colon cancer 22 AAH33268  
 Human normal bladd 20 AAZ42164  
 Human colon cancer 21 AAC98058  
 Human colon cancer 21 AAZ80570  
 Human cDNA encodin 24 ABK35125  
 Human colon cancer 24 ABOS893  
 Human foetal liver 22 ABAS1954  
 Probe #238 for gen 22 ABP21772  
 Human brain expres 21 AAU0240  
 Human bone marrow 22 AAU25682  
 Probe #242 for gen 22 AAU10309  
 Probe #248 used to 22 AAU131562  
 Probe #246 used to 22 AAU100245  
 Human liver single 23 ABS25267  
 Human genome-deriv 24 ABU0260  
 Nucleotide sequenc 24 ABUS9087  
 Human secretary pr 24 AAU54582  
 Human secreted pro 24 AAU25601  
 Frog embryonic gen 24 ABS77417  
 Human colon cancer 23 AAH35088  
 Frog embryonic gen 24 ABS77066  
 Human breast cell 22 ABAB6733  
 Human foetal liver 22 ABAB64613  
 Probe #10206 for g 22 AAK31740  
 Human brain expres 22 AAK13051  
 Human bone marrow 22 AAK31781  
 Probe #13469 used 22 AAU119587  
 Probe #5298 used t 22 AAU144783  
 Human liver single 22 AAU105307  
 Human genome-deriv 23 ABS383864  
 Human brain Express 24 ABS12860  
 Expressed Sequence 24 AQS9054  
 FISH clone, BK260 24 AQS9642  
 Datasheet : 24 AAV69142  
 45 307  
 44 342  
 46 342  
 47 342  
 48 342  
 49 342  
 50 342  
 51 342  
 52 342  
 53 342  
 54 342  
 55 342  
 56 342  
 57 342  
 58 342  
 59 342  
 60 342  
 61 342  
 62 342  
 63 342  
 64 342  
 65 342  
 66 342  
 67 342  
 68 342  
 69 342  
 70 342  
 71 342  
 72 342  
 73 342  
 74 342  
 75 342  
 76 342  
 77 342  
 78 342  
 79 342  
 80 342  
 81 342  
 82 342  
 83 342  
 84 342  
 85 342  
 86 342  
 87 342  
 88 342  
 89 342  
 90 342  
 91 342  
 92 342  
 93 342  
 94 342  
 95 342  
 96 342  
 97 342  
 98 342  
 99 342  
 100 342  
 101 342  
 102 342  
 103 342  
 104 342  
 105 342  
 106 342  
 107 342  
 108 342  
 109 342  
 110 342  
 111 342  
 112 342  
 113 342  
 114 342  
 115 342  
 116 342  
 117 342  
 118 342  
 119 342  
 120 342  
 121 342  
 122 342  
 123 342  
 124 342  
 125 342  
 126 342  
 127 342  
 128 342  
 129 342  
 130 342  
 131 342  
 132 342  
 133 342  
 134 342  
 135 342  
 136 342  
 137 342  
 138 342  
 139 342  
 140 342  
 141 342  
 142 342  
 143 342  
 144 342  
 145 342  
 146 342  
 147 342  
 148 342  
 149 342  
 150 342  
 151 342  
 152 342  
 153 342  
 154 342  
 155 342  
 156 342  
 157 342  
 158 342  
 159 342  
 160 342  
 161 342  
 162 342  
 163 342  
 164 342  
 165 342  
 166 342  
 167 342  
 168 342  
 169 342  
 170 342  
 171 342  
 172 342  
 173 342  
 174 342  
 175 342  
 176 342  
 177 342  
 178 342  
 179 342  
 180 342  
 181 342  
 182 342  
 183 342  
 184 342  
 185 342  
 186 342  
 187 342  
 188 342  
 189 342  
 190 342  
 191 342  
 192 342  
 193 342  
 194 342  
 195 342  
 196 342  
 197 342  
 198 342  
 199 342  
 200 342  
 201 342  
 202 342  
 203 342  
 204 342  
 205 342  
 206 342  
 207 342  
 208 342  
 209 342  
 210 342  
 211 342  
 212 342  
 213 342  
 214 342  
 215 342  
 216 342  
 217 342  
 218 342  
 219 342  
 220 342  
 221 342  
 222 342  
 223 342  
 224 342  
 225 342  
 226 342  
 227 342  
 228 342  
 229 342  
 230 342  
 231 342  
 232 342  
 233 342  
 234 342  
 235 342  
 236 342  
 237 342  
 238 342  
 239 342  
 240 342  
 241 342  
 242 342  
 243 342  
 244 342  
 245 342  
 246 342  
 247 342  
 248 342  
 249 342  
 250 342  
 251 342  
 252 342  
 253 342  
 254 342  
 255 342  
 256 342  
 257 342  
 258 342  
 259 342  
 260 342  
 261 342  
 262 342  
 263 342  
 264 342  
 265 342  
 266 342  
 267 342  
 268 342  
 269 342  
 270 342  
 271 342  
 272 342  
 273 342  
 274 342  
 275 342  
 276 342  
 277 342  
 278 342  
 279 342  
 280 342  
 281 342  
 282 342  
 283 342  
 284 342  
 285 342  
 286 342  
 287 342  
 288 342  
 289 342  
 290 342  
 291 342  
 292 342  
 293 342  
 294 342  
 295 342  
 296 342  
 297 342  
 298 342  
 299 342  
 300 342  
 301 342  
 302 342  
 303 342  
 304 342  
 305 342  
 306 342  
 307 342  
 308 342  
 309 342  
 310 342  
 311 342  
 312 342  
 313 342  
 314 342  
 315 342  
 316 342  
 317 342  
 318 342  
 319 342  
 320 342  
 321 342  
 322 342  
 323 342  
 324 342  
 325 342  
 326 342  
 327 342  
 328 342  
 329 342  
 330 342  
 331 342  
 332 342  
 333 342  
 334 342  
 335 342  
 336 342  
 337 342  
 338 342  
 339 342  
 340 342  
 341 342  
 342 342  
 343 342  
 344 342  
 345 342  
 346 342  
 347 342  
 348 342  
 349 342  
 350 342  
 351 342  
 352 342  
 353 342  
 354 342  
 355 342  
 356 342  
 357 342  
 358 342  
 359 342  
 360 342  
 361 342  
 362 342  
 363 342  
 364 342  
 365 342  
 366 342  
 367 342  
 368 342  
 369 342  
 370 342  
 371 342  
 372 342  
 373 342  
 374 342  
 375 342  
 376 342  
 377 342  
 378 342  
 379 342  
 380 342  
 381 342  
 382 342  
 383 342  
 384 342  
 385 342  
 386 342  
 387 342  
 388 342  
 389 342  
 390 342  
 391 342  
 392 342  
 393 342  
 394 342  
 395 342  
 396 342  
 397 342  
 398 342  
 399 342  
 400 342  
 401 342  
 402 342  
 403 342  
 404 342  
 405 342  
 406 342  
 407 342  
 408 342  
 409 342  
 410 342  
 411 342  
 412 342  
 413 342  
 414 342  
 415 342  
 416 342  
 417 342  
 418 342  
 419 342  
 420 342  
 421 342  
 422 342  
 423 342  
 424 342  
 425 342  
 426 342  
 427 342  
 428 342  
 429 342  
 430 342  
 431 342  
 432 342  
 433 342  
 434 342  
 435 342  
 436 342  
 437 342  
 438 342  
 439 342  
 440 342  
 441 342  
 442 342  
 443 342  
 444 342  
 445 342  
 446 342  
 447 342  
 448 342  
 449 342  
 450 342  
 451 342  
 452 342  
 453 342  
 454 342  
 455 342  
 456 342  
 457 342  
 458 342  
 459 342  
 460 342  
 461 342  
 462 342  
 463 342  
 464 342  
 465 342  
 466 342  
 467 342  
 468 342  
 469 342  
 470 342  
 471 342  
 472 342  
 473 342  
 474 342  
 475 342  
 476 342  
 477 342  
 478 342  
 479 342  
 480 342  
 481 342  
 482 342  
 483 342  
 484 342  
 485 342  
 486 342  
 487 342  
 488 342  
 489 342  
 490 342  
 491 342  
 492 342  
 493 342  
 494 342  
 495 342  
 496 342  
 497 342  
 498 342  
 499 342  
 500 342  
 501 342  
 502 342  
 503 342  
 504 342  
 505 342  
 506 342  
 507 342  
 508 342  
 509 342  
 510 342  
 511 342  
 512 342  
 513 342  
 514 342  
 515 342  
 516 342  
 517 342  
 518 342  
 519 342  
 520 342  
 521 342  
 522 342  
 523 342  
 524 342  
 525 342  
 526 342  
 527 342  
 528 342  
 529 342  
 530 342  
 531 342  
 532 342  
 533 342  
 534 342  
 535 342  
 536 342  
 537 342  
 538 342  
 539 342  
 540 342  
 541 342  
 542 342  
 543 342  
 544 342  
 545 342  
 546 342  
 547 342  
 548 342  
 549 342  
 550 342  
 551 342  
 552 342  
 553 342  
 554 342  
 555 342  
 556 342  
 557 342  
 558 342  
 559 342  
 560 342  
 561 342  
 562 342  
 563 342  
 564 342  
 565 342  
 566 342  
 567 342  
 568 342  
 569 342  
 570 342  
 571 342  
 572 342  
 573 342  
 574 342  
 575 342  
 576 342  
 577 342  
 578 342  
 579 342  
 580 342  
 581 342  
 582 342  
 583 342  
 584 342  
 585 342  
 586 342  
 587 342  
 588 342  
 589 342  
 590 342  
 591 342  
 592 342  
 593 342  
 594 342  
 595 342  
 596 342  
 597 342  
 598 342  
 599 342  
 600 342  
 601 342  
 602 342  
 603 342  
 604 342  
 605 342  
 606 342  
 607 342  
 608 342  
 609 342  
 610 342  
 611 342  
 612 342  
 613 342  
 614 342  
 615 342  
 616 342  
 617 342  
 618 342  
 619 342  
 620 342  
 621 342  
 622 342  
 623 342  
 624 342  
 625 342  
 626 342  
 627 342  
 628 342  
 629 342  
 630 342  
 631 342  
 632 342  
 633 342  
 634 342  
 635 342  
 636 342  
 637 342  
 638 342  
 639 342  
 640 342  
 641 342  
 642 342  
 643 342  
 644 342  
 645 342  
 646 342  
 647 342  
 648 342  
 649 342  
 650 342  
 651 342  
 652 342  
 653 342  
 654 342  
 655 342  
 656 342  
 657 342  
 658 342  
 659 342  
 660 342  
 661 342  
 662 342  
 663 342  
 664 342  
 665 342  
 666 342  
 667 342  
 668 342  
 669 342  
 670 342  
 671 342  
 672 342  
 673 342  
 674 342  
 675 342  
 676 342  
 677 342  
 678 342  
 679 342  
 680 342  
 681 342  
 682 342  
 683 342  
 684 342  
 685 342  
 686 342  
 687 342  
 688 342  
 689 342  
 690 342  
 691 342  
 692 342  
 693 342  
 694 342  
 695 342  
 696 342  
 697 342  
 698 342  
 699 342  
 700 342  
 701 342  
 702 342  
 703 342  
 704 342  
 705 342  
 706 342  
 707 342  
 708 342  
 709 342  
 710 342  
 711 342  
 712 342  
 713 342  
 714 342  
 715 342  
 716 342  
 717 342  
 718 342  
 719 342  
 720 342  
 721 342  
 722 342  
 723 342  
 724 342  
 725 342  
 726 342  
 727 342  
 728 342  
 729 342  
 730 342  
 731 342  
 732 342  
 733 342  
 734 342  
 735 342  
 736 342  
 737 342  
 738 342  
 739 342  
 740 342  
 741 342  
 742 342  
 743 342  
 744 342  
 745 342  
 746 342  
 747 342  
 748 342  
 749 342  
 750 342  
 751 342  
 752 342  
 753 342  
 754 342  
 755 342  
 756 342  
 757 342  
 758 342  
 759 342  
 760 342  
 761 342  
 762 342  
 763 342  
 764 342  
 765 342  
 766 342  
 767 342  
 768 342  
 769 342  
 770 342  
 771 342  
 772 342  
 773 342  
 774 342  
 775 342  
 776 342  
 777 342  
 778 342  
 779 342  
 780 342  
 781 342  
 782 342  
 783 342  
 784 342  
 785 342  
 786 342  
 787 342  
 788 342  
 789 342  
 790 342  
 791 342  
 792 342  
 793 342  
 794 342  
 795 342  
 796 342  
 797 342  
 798 342  
 799 342  
 800 342  
 801 342  
 802 342  
 803 342  
 804 342  
 805 342  
 806 342  
 807 342  
 808 342  
 809 342  
 810 342  
 811 342  
 812 342  
 813 342  
 814 342  
 815 342  
 816 342  
 817 342  
 818 342  
 819 342  
 820 342  
 821 342  
 822 342  
 823 342  
 824 342  
 825 342  
 826 342  
 827 342  
 828 342  
 829 342  
 830 342  
 831 342  
 832 342  
 833 342  
 834 342  
 835 342  
 836 342  
 837 342  
 838 342  
 839 342  
 840 342  
 841 342  
 842 342  
 843 342  
 844 342  
 845 342  
 846 342  
 847 342  
 848 342  
 849 342  
 850 342  
 851 342  
 852 342  
 853 342  
 854 342  
 855 342  
 856 342  
 857 342  
 858 342  
 859 342  
 860 342  
 861 342  
 862 342  
 863 342  
 864 342  
 865 342  
 866 342  
 867 342  
 868 342  
 869 342  
 870 342  
 871 342  
 872 342  
 873 342  
 874 342  
 875 342  
 876 342  
 877 342  
 878 342  
 879 342  
 880 342  
 881 342  
 882 342  
 883 342  
 884 342  
 885 342  
 886 342  
 887 342  
 888 342  
 889 342  
 890 342  
 891 342  
 892 342  
 893 342  
 894 342  
 895 342  
 896 342  
 897 342  
 898 342  
 899 342  
 900 342  
 901 342  
 902 342  
 903 342  
 904 342  
 905 342  
 906 342  
 907 342  
 908 342  
 909 342  
 910 342  
 911 342  
 912 342  
 913 342  
 914 342  
 915 342  
 916 342  
 917 342  
 918 342  
 919 342  
 920 342  
 921 342  
 922 342  
 923 342  
 924 342  
 925 342  
 926 342  
 927 342  
 928 342  
 929 342  
 930 342  
 931 342  
 932 342  
 933 342  
 934 342  
 935 342  
 936 342  
 937 342  
 938 342  
 939 342  
 940 342  
 941 342  
 942 342  
 943 342  
 944 342  
 945 342  
 946 342  
 947 342  
 948 342  
 949 342  
 950 342  
 951 342  
 952 342  
 953 342  
 954 342  
 955 342  
 956 342  
 957 342  
 958 342  
 959 342  
 960 342  
 961 342  
 962 342  
 963 342  
 964 342  
 965 342  
 966 342  
 967 342  
 968 342  
 969 342  
 970 342  
 971 342  
 972 342  
 973 342  
 974 342  
 975 342  
 976 342  
 977 342  
 978 342  
 979 342  
 980 342  
 981 342  
 982 342  
 983 342  
 984 342  
 985 342  
 986 342  
 987 342  
 988 342  
 989 342  
 990 342  
 991 342  
 992 342  
 993 342  
 994 342  
 995 342  
 996 342  
 997 342  
 998 342  
 999 342  
 1000 342

Page 2



TELEFAX:	650-845-4166	
TELEX:	<Unknown>	
INFORMATION FOR SEQ ID NO: 2:		
SEQUENCE CHARACTERISTICS:		
LENGTH:	2079 base pairs	
TYPE: nucleic acid		
STRANDEDNESS: single		
TOPOLOGY: linear		
IMMEDIATE SOURCE:		
LIBRARY: TIYMNOTOS		
CLONE: 2805795		
SEQUENCE DESCRIPTION: SEQ ID NO: 2:		
Query Match	100.0%	Score 2079; DB 9; Length 2079;
Best Local Similarity	100.0%	Pred. No. 0;
Matches	2079;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	61	CCAGCGCGTGAAGCTGGCGAACCGGGAGCGCGCTCGGAGGGAGTGGAACCTGGANTGGCTTTGCTG 60
Db	61	CCAGCGCGTGAAGCTGGCGAACCGGGAGCGCGCTCGGAGGGAGTGGAACCTGGANTGGCTTTGCTG 60
Qy	121	GCAGCGAGCAGCGTTTCGGCCAAACCCGATAACCATCGGAGAGGATTTCGGCTCGGCCAACCGGA 120
Db	121	GCAGCGAGCAGCGTTTCGGCCAAACCCGATAACCATCGGAGAGGATTTCGGCTCGGCCAACCGGA 120
Qy	1201	TATGCCAAGGTAAAGCTGGCTGTTGCAAGGGCTTAACAGGCTTACGTGAACTGGCTGATCTGGGT 1200
Db	1201	TATGCCAAGGTAAAGCTGGCTGTTGCAAGGGCTTAACAGGCTTACGTGAACTGGCTGATCTGGGT 1200
Qy	1081	TGAAAGATCACAAAACAGTGCAGGCCAGGGTAGTTAATGAGTTAAAGTCAGTGTAAAGTCAGGT 1140
Db	1081	TGAAAGATCACAAAACAGTGCAGGCCAGGGTAGTTAATGAGTTAAAGTCAGTGTAAAGTCAGGT 1140
Qy	1021	CATTATAACCACACGGATGACGAAAAAACAGAGCTTNGCTTCTTGATAA 1080
Db	1021	CATTATAACCACACGGATGACGAAAAAACAGAGCTTNGCTTCTTGATAA 1080
Qy	1141	CTGGGGAAATGTGGAACTGTGAACTGGCTGATCTGGATCTAGAGATCTGTGATCTGGGT 1260
Db	1141	CTGGGGAAATGTGGAACTGTGAACTGGCTGATCTGGATCTAGAGATCTGTGATCTGGGT 1260
Qy	1201	TATGCCAAGGTAAAGCTGGCTGTTGCAAGGGCTTAACAGGCTTACGTGAACTGGCTGATCTGGGT 1260
Db	1201	TATGCCAAGGTAAAGCTGGCTGTTGCAAGGGCTTAACAGGCTTACGTGAACTGGCTGATCTGGGT 1260
Qy	1261	TTTAGAAAGGCCATTAGTCAGTTGGGAAACTCTGAACAGGTGAGGTAAAGGATTA 1320
Db	1261	TTTAGAAAGGCCATTAGTCAGTTGGGAAACTCTGAACAGGTGAGGTAAAGGATTA 1320
Qy	1321	TGCGTTCATTCATTGTGAGGGAGATGGTGTCAAGGGTATGGAGAAATGATGG 1380
Db	1321	TGCGTTCATTCATTGTGAGGGAGATGGTGTCAAGGGTATGGAGAAATGATGG 1380
Qy	1381	CAAGACTTGGAGGGAGAAATTGGAAATTGGTGTGGCTGAGGTGAGGTGGATGATGG 1440
Db	1381	CAAGACTTGGAGGGAGAAATTGGAAATTGGTGTGGCTGAGGTGAGGTGGATGATGG 1440
Qy	1441	GAAGAAAGAAAAGCTCAGGGAAAGCAGCAAAATAAATGTAAGGATTACTACTA 1500
Db	1441	GAAGAAAGAAAAGCTCAGGGAAAGCAGCAAAATAAATGTAAGGATTACTACTA 1500
Qy	1501	TTATGTCACCTCATATGCCCTCCACAAAGGTGAGGTGGATGGTGG 1560
Db	1501	TTATGTCACCTCATATGCCCTCCACAAAGGTGAGGTGGATGGTGG 1560
Qy	1561	TTATGATPATCCMCAGATTATGGATTATGGATTATTATGGTATG 1620
Db	1561	TTATGATPATCCMCAGATTATGGATTATGGATTATGGTATG 1620
Qy	1621	TTACCATAACTACTGTGGTGGPATGAGATCCATACTATGGTTAGGATTTCAGT 1680
Db	1621	TTACCATAACTACTGTGGTGGPATGAGATCCATACTATGGTTAGGATTTCAGT 1680
Qy	1681	TGGAGCTTACAGGGAGGGGGTGAAGGAGCAACGGTCTCAGAGCTTGGGG 1740
Db	1681	TGGAGCTTACAGGGAGGGGGTGAAGGAGCAACGGTCTCAGAGCTTGGGG 1740
Qy	1741	GGCTGCTCTCCCTCCCGGGTTAGGCCGGTTATTCACAGAGGGGTCTGGATCAGGCTGTGG 1800
Db	1741	GGCTGCTCTCCCTCCCGGGTTAGGCCGGTTATTCACAGAGGGGTCTGGATCAGGCTGTGG 1800
Qy	1801	AGGCCTTCAGCAGGGAAAGGGCTGAGGGCTCTGACCTGTACATGAGACTGA 1860
Db	1801	AGGCCTTCAGCAGGGAAAGGGCTGAGGGCTCTGACCTGTACATGAGACTGA 1860
Qy	1861	CTTGCTTATCTGGGATTAACCCAGAGCTGGTCACTGGTTTGTACAA 1920
Db	1861	CTTGCTTATCTGGGATTAACCCAGAGCTGGTCACTGGTTTGTACAA 1920

Result No.	Score	Query Match Length DB ID	Description
1	2079	100.0	US-09-978-242-2 Sequence 2, App1
2	1844	88.7	Sequence 475, App
3	1826.2	87.8	Sequence 3, App1
4	1813.4	87.2	Sequence 1042, App
5	882	42.4	Sequence 127, App
6	882	42.4	Sequence 85, App1
c	7	871.6	Sequence 199, App
8	839.8	40.4	Sequence 1381, App
9	833.6	40.1	Sequence 334, App
10	730	35.1	Sequence 68, App1
11	730	35.1	Sequence 1024, App1
12	498.8	24.0	Sequence 654, App
13	487	23.4	Sequence 263, App
14	434.8	20.9	Sequence 9841, App
15	425	20.4	Sequence 238, App

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length DB ID	Description
1	2079	100.0	US-09-978-242-2 Sequence 2, App1
2	1844	88.7	Sequence 475, App
3	1826.2	87.8	Sequence 3, App1
4	1813.4	87.2	Sequence 1042, App
5	882	42.4	Sequence 127, App
6	882	42.4	Sequence 85, App1
c	7	871.6	Sequence 199, App
8	839.8	40.4	Sequence 1381, App
9	833.6	40.1	Sequence 334, App
10	730	35.1	Sequence 68, App1
11	730	35.1	Sequence 1024, App1
12	498.8	24.0	Sequence 654, App
13	487	23.4	Sequence 263, App
14	434.8	20.9	Sequence 9841, App
15	425	20.4	Sequence 238, App

Scoring table: IDENTITY\_NUC Gapop 10.0 , Gapext 1.0  
 Searched: 2356869 seqs, 1788235258 residues  
 Total number of hits satisfying chosen parameters: 4713738  
 Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries  
 Database : Published Applications NA:  
 1: /cgn2\_6/podata/1/pubnpna/us07\_pubcomb.seq;  
 2: /cgn2\_6/podata/1/pubnpna/us07\_pubcomb.seq;  
 3: /cgn2\_6/podata/1/pubnpna/us06\_pubseq.\*  
 4: /cgn2\_6/podata/1/pubnpna/us06\_pubcomb.seq;  
 5: /cgn2\_6/podata/1/pubnpna/us07\_pubseq.\*  
 6: /cgn2\_6/podata/1/pubnpna/pctus\_pubcomb.seq.\*  
 7: /cgn2\_6/podata/1/pubnpna/us08\_pubseq.\*  
 8: /cgn2\_6/podata/1/pubnpna/us08\_pubcomb.seq.\*  
 9: /cgn2\_6/podata/1/pubnpna/us09\_pubcomb.seq.\*  
 10: /cgn2\_6/podata/1/pubnpna/us09b\_pubcomb.seq.\*  
 11: /cgn2\_6/podata/1/pubnpna/us09c\_pubcomb.seq.\*  
 12: /cgn2\_6/podata/1/pubnpna/us09\_pubseq.\*  
 13: /cgn2\_6/podata/1/pubnpna/us09\_pubseq.\*  
 14: /cgn2\_6/podata/1/pubnpna/us10\_pubcomb.seq.\*  
 15: /cgn2\_6/podata/1/pubnpna/us10b\_pubcomb.seq.\*  
 16: /cgn2\_6/podata/1/pubnpna/us10\_pubseq.\*  
 17: /cgn2\_6/podata/1/pubnpna/us10\_pubseq.\*  
 18: /cgn2\_6/podata/1/pubnpna/us60\_pubcomb.seq.\*  
 RESULT 1  
 US-09-978-242-2  
 ; Sequence 2, Application US/09978242  
 ; Patent No. US20020098566A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; Yue, Henry  
 ; Corley, Neil C.  
 ; Shah, Purvi  
 ; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/978,242  
 FILING DATE: 15-Oct-2001  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 09/241,333  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: 08/990,114  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0451 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555

## ALIGNMENTS

US-09-978-242-2  
 ; Sequence 2, Application US/09978242  
 ; Patent No. US20020098566A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; Yue, Henry  
 ; Corley, Neil C.  
 ; Shah, Purvi  
 ; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/978,242  
 FILING DATE: 15-Oct-2001  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 09/241,333  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: 08/990,114  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0451 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555

REFERENCE	3	Shibata,K., Itoh,M., Aizawa,K., Nagaoa,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Kitaunai,T., Tashiro,H., Itoh,M., Nishi,K., Kiyosawa,M., Yashima,T., Hazama,M., Nishine,T., Harada,A., Sumi,N., Ishii,Y., Nakamura,S., Hidemoto,O., Sakai,C., Ikeyama,T., Kashiwagi,K., Yamamoto,R., Matsumoto,T., Togawa,Y., Izawa,M., Ohara,E., Watanuki,M., Fujimoto,S., Inoue,K., Togawa,Y., Ozawa,K., Tanaka,T., Matsuru,S., Kawai,J., Yoneda,Y., Ishikawa,T., Okuno,H., Matsuru,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.	SOURCE	1. .3838 organism="Mus musculus" /mol type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:9430046J23" /db_xref="ITAXON:10090" /clone="9430046J23" /tissue_type="embryonic body between diaphragm region and neck" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="12 days embryo" /note="unnamed protein product; NS1-associated protein 1 (MGI:MGI:1891650, GB AB055725, evidence: BLASTN, 99%, match=31.07) putative" /codon_start=1 /protein_id="BAC28852.1" /db_xref="GI: 26330244" /translation="MATEHYUNGTEPEMDTSAVHSNENPQTLIDAGLPOKVAEKLDBIYAGIVAHSDLDRAEALKEPNEDGALAVLQOQFKDSUHVONKSAPLGWVKT ROREKGTYVADSSKGPDDEAKIKALLERTGYTLDVTGQRQYGGPPDSDYSGQPSV GTEIFVGKIPRDFFEDVPLPKAGPNDPLTGLNRGTAUTPCTEKAQAE AVKLNNHETRSRGHTGVCSVANNRLFVEISPKSTKEOLLEPSTKVTEBLDVLY HOPDDEKKRNKGFCFLYEIDHDKTAQARRRUMSGKVWGMNGTVWMAPIBDPDEVM AKVFLVFLNLTNTILEKSFQGKLERVKLKDAPTHPDERGAVKAMEM GRDLIGENELTIVFAKEPDKRKERKAQKANQMDYYGGPHMPPTGRGRGG RGGYGYPPDYGYEDYYGYDYYNYRGYEDPYGYEDFOVGARGRRGGARGAAP SRGRGAAPPRGRAGSQRGGPSARGVARGGAAQQRGCGQKGVEAGGPDLLO"
AUTHORS	4	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Iwaza,M., Nishi,K., Kiyosawa,M., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,Y., Tomita,M., Kadota,K., Matsuda,H., Ashburner,M., Battalov,S., Cesavant,T., Fleischmann,W., Gaasterland,T., Gissel,A., King,B., Kochiwa,H., Kiehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pecole,G., Quackenbush,J., Schriml,L.M., Straubli,F., Suzuki,R., Tomita,M., Wagner,B., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Boujunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Garibotti,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriiguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seye,T., Shiba,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S., and Hayashizaki,Y.	CDS	167. .1855 /note="sequencing pipeline with 384 multicapillary sequencer sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 1.0766861
AUTHORS	5	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Iwaza,M., Nishi,K., Kiyosawa,M., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,Y., Tomita,M., Kadota,K., Matsuda,H., Ashburner,M., Battalov,S., Cesavant,T., Fleischmann,W., Gaasterland,T., Gissel,A., King,B., Kochiwa,H., Kiehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pecole,G., Quackenbush,J., Schriml,L.M., Straubli,F., Suzuki,R., Tomita,M., Wagner,B., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Boujunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Garibotti,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriiguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seye,T., Shiba,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S., and Hayashizaki,Y.	CDNA	Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
AUTHORS	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	REFERENCE	Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
AUTHORS	6	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	JOURNAL	5
AUTHORS	6	Nature 420, 563-573 (2002)	REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS	7	Adachi,J., Aizawa,K., Akimura,T., Arikawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Horii,F., Imorani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katsch,H., Kawai,J., Koijima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Morata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sogabe,Y., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Tagawa,A., Takeda,Y., Tanaka,T., Tomaru,A., Toyoda,T., Yasunishi,A., Muramatsu,M., and Hayashizaki,Y.	JOURNAL	6
AUTHORS	7	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gbc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/), Tel: 81-45-503-9222, Fax: 81-45-503-9216).	REFERENCE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to
AUTHORS	8	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://fantom.gsc.riken.go.jp/	JOURNAL	8
COMMENT	9	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://fantom.gsc.riken.go.jp/	JOURNAL	9
FEATURES	10	578 GCACCTTGAAAAGAACGGCTACACTGTGACATAGTTAGTAGAAGAGCTATTGAGCTTTA 577 GAAAAACAGGGACCAAGAACGGCTACACTGTGACATAGTTAGTAGAAGAGCTATTGAGCTTTA 576 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 575 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 574 TCTCTGTTCAAGAACGGCTACACTGTGACATAGTTAGTAGAAGAGCTATTGAGCTTTA 573 GAAACACAGGGACCAAGAACGGCTACACTGTGACATAGTTAGTAGAAGAGCTATTGAGCTTTA 572 GAAACACAGGGACCAAGAACGGCTACACTGTGACATAGTTAGTAGAAGAGCTATTGAGCTTTA 571 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 570 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 569 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 568 TCTCTGTTCAAGAACGGCTACACTGTGACATAGTTAGTAGAAGAGCTATTGAGCTTTA 567 GAAACACAGGGACCAAGAACGGCTACACTGTGACATAGTTAGTAGAAGAGCTATTGAGCTTTA 566 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 565 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 564 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 563 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 562 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 561 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 560 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 559 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 558 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 557 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 556 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 555 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 554 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 553 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 552 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 551 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 550 AAAGATTCAATGAAAGCCGTCAGTGCATGGAGCTTCAACAGTTAACAGCTATGCTACAGTTAACATTAG 549 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 548 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 547 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 546 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 545 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 544 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 543 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 542 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 541 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 540 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 539 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 538 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 537 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 536 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 535 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 534 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 533 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 532 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 531 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 530 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 529 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 528 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 527 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 526 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 525 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 524 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 523 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 522 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 521 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 520 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 519 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 518 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 517 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 516 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 515 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 514 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 513 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 512 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 511 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 510 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 509 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 508 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 507 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 506 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 505 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 504 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 503 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 502 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 501 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 500 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 499 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 498 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 497 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 496 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 495 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 494 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 493 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 492 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 491 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 490 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 489 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 488 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 487 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 486 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 485 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 484 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 483 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 482 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 481 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 480 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 479 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 478 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 477 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 476 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 475 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 474 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 473 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 472 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 471 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 470 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 469 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 468 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 467 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 466 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 465 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 464 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 463 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 462 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 461 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 460 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 459 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 458 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 457 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 456 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 455 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 454 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 453 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 452 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 451 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 450 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 449 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 448 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 447 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 446 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 445 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 444 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 443 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 442 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 441 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 440 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 439 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 438 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 437 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 436 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 435 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 434 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 433 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 432 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 431 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 430 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 429 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 428 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 427 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 426 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 425 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 424 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 423 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 422 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 421 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 420 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 419 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 418 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 417 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 416 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 415 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 414 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 413 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 412 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 411 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 410 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 409 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 408 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 407 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 406 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 405 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 404 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 403 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 402 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 401 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 400 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 399 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 398 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 397 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 396 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 395 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 394 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 393 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 392 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 391 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 390 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 389 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 388 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 387 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 386 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 385 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 384 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 383 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 382 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 381 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 380 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 379 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 378 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 377 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 376 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 375 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 374 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 373 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 372 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 371 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 370 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 369 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 368 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 367 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 366 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 365 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 364 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 363 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAGAGA 362 ATTACGTTGAGGCTAGTGCACATAGTTAGTAGAAGAGCTATTGAGCTTTA 361 GCTCATGTTGAGAACAAGTGGCTTATGTCAGTGCATGGAGCTTACAGGAG		

Result No.	Score	Query	Match length	DB ID	Description	SUMMARIES
1	1792.4	gb_gss1*	89.0	AK034845	AK034845 Mus musculus	1
2	1792.4	gb_gss1*	86.2	2141	AK076026 Mus musculus	2
3	1787.4	gb_gss1*	86.0	2030	AK077588 Mus musculus	3
4	1537	gb_gss1*	73.9	4064	AK0833398 Mus musculus	4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Run on:	January 29, 2004, 00:06:34 ; Search time 2986 Seconds (without alignments)	5	1494.8	AK078158 Mus musculus	3289 11
Title:	US-09-978-242-2	6	1386.6	BC021932 Homo sapi	3289 11
Perfect score:	2079	7	1353.6	BC024283 Homo sapi	3289 11
Sequence:	1 GGGCGCCGCGCACGGG,..... GTTATAACTAAAGCTACT 2079	8	1343.6	BC019360 Homo sapi	3289 11
Scoring table:	IDENTITY_NUC Gapop 10.0 , Gapext 1.0	9	1331.6	BC009176 Homo sapi	3289 11
Searched:	22781392 seqs, 12152238056 residues	10	1327.4	BC015575 Homo sapi	3289 11
Total number of hits satisfying chosen parameters:	45562784	11	983.8	AL541211 AL541211	3289 11
Minimum DB seq length:	0	12	975.2	AL541223 AL541223	3289 11
Maximum DB seq length:	2000000000	13	921.6	AL541248 AL541248	3289 11
Post-processing:	Minimum Match 0% Maximum Match 100%	14	884.2	BC026850 Mus musculus	3289 11
Database :	EST:*	15	883	BN466029 AGENCOURT	3289 11
	1: em_estba:*	16	866.6	B0686098 AGENCOURT	3289 11
	2: em_estchum:*	17	862.8	BM809311 AGENCOURT	3289 11
	3: em_estin:*	18	844.6	BN945048 AGENCOURT	3289 11
	4: em_estmu:*	19	838.2	BX420656 BX420656	3289 11
	5: em_estov:*	20	836.6	BU056813 AGENCOURT	3289 11
	6: em_estpl:*	21	831.8	BU194176 AGENCOURT	3289 11
	7: em_estro:*	22	829.2	BUT83896 AGENCOURT	3289 11
	8: em_ntc:*	23	826.8	BM451217 AGENCOURT	3289 11
	9: gb_est1:*	24	807.6	BU845727 AGENCOURT	3289 11
	10: gb_est2:*	25	768	BM43159 AGENCOURT	3289 11
	11: gb_htc:*	26	757.8	BU056813 AGENCOURT	3289 11
	12: gb_est3:*	27	756.6	BU194176 AGENCOURT	3289 11
	13: gb_est4:*	28	747.6	BU731663 601567050	3289 11
	14: gb_est5:*	29	744.6	BU650432 AGENCOURT	3289 11
	15: em_eston:*	30	740.8	BU680994 AGENCOURT	3289 11
	16: em_eston:*	31	740.2	BU950249 AGENCOURT	3289 11
	17: em_gss_hum:*	32	736.2	BU959242 AGENCOURT	3289 11
	18: em_gss_inv:*	33	729	AV298691 AV298691	3289 11
	19: em_gss_pln:*	34	728.6	CB182505 AGENCOURT	3289 11
	20: em_gss_rtc:*	35	721.2	BB731663 601567050	3289 11
	21: em_gss_mam:*	36	719	BB43866 UI-M-FCO-	3289 11
	22: em_gss_nus:*	37	715.6	BB407142 601301856	3289 11
	23: em_gss_pro:*	38	712.8	BB026689 602293533	3289 11
	24: em_gss_rod:*	39	710.6	BB158562 602403281	3289 11
	25: em_gss_phg:*	40	709.8	BB282140 602403281	3289 11
	26: em_gss_vrl:*	41	704.6	BB571781 UI-M-FCO-	3289 11
	27: em_gss_vrt:*	42	700.4	BB1661316 603346285	3289 11
	28: gb_gss1:*	43	700.2	BBE60849 601346126	3289 11
	29: gb_gss2:*	44	698.8	BB558142 601240061	3289 11
		45	697.8	BB191822 AGENCOURT	3289 11

## ALIGNMENTS

RESULT 1	AK034845 LOCUS	AK034845 Mus musculus 12 days embryo linear mRNA
	DEFINITION	Mus musculus 12 days embryo linear mRNA and neck cDNA, RIKEN full-length enriched library, clone: 930046223 product: NS1-associated protein 1, full insert sequence.
	ACCESSION	AK034845
	VERSION	.1 GI:26330243
	KEYWORDS	HTC; CAP trapper.
	SOURCE	Mus musculus (house mouse)
	ORGANISM	Mus musculus
REFERENCE	AUTHORS	Carninci, P. and Hayashizaki, Y.
	TITLE	High-efficiency full-length cDNA cloning
	JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
	MEDLINE	99229253
	PUBMED	10349636
REFERENCE 2	AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibusawa, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
	TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
	JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
	MEDLINE	2049374
	PUBMED	11042159